

A COGNITIVE APPROACH TO
UNDERSTANDING THE BEHAVIOUR
OF MUSEUM VISITORS

UNIVERSITY OF LONDON
INSTITUTE OF EDUCATION LIBRARY
NOT TO BE REMOVED FROM LIBRARY

M. B. ALT

University of London Institute of Education
Thesis submitted for the Degree of PhD 1982

ACKNOWLEDGEMENTS

I am grateful to Dr. R. S. Miles, Head of the Department of Public Services, at the British Museum (Natural History) for his encouragement throughout and to Dr. Michael Hills and Ms. Kay Shaw of the Biometrics section at the Museum for their helpful advice on statistical matters and methodology.

Thanks are also due to Mr. Alex Gray in the Museum's electronic workshop for translating my ideas about methods of studying visitors' behaviour into a working system, to the assistants who helped me collect the data and to all the other members of the Museum's staff who helped in one way or another to make the project possible in the hurly-burly of an institution that receives three million visitors a year.

Special thanks are offered to Dr. Norman Worrall at the Institute of Education who pinned down my more fanciful ideas and who has tried his best to turn my abstruse prose-style into a clear exposition.

Needless to say, I am grateful to the thousands of unsuspecting visitors who generously gave their time over to interviews when they came to look at exhibits.

Last but by no means least I should like to thank my wife Lesley who took over all the duties of looking after a young family to allow me to write this thesis. Without her help and insistence it would never have been written.

ABSTRACT

Previous studies of the behaviour of museum visitors and the theories of exhibit effectiveness that have stemmed from them are reviewed and criticised on a number of grounds. The principle criticism is directed at the prevailing notion of the Museum Visitor as a passive being whose actions are determined by internal or external forces over which he or she has no control. A new theory of exhibit effectiveness is presented that accepts the visitor as an active agent who acts in accordance with his or her perceptions of individual displays. Real-world exhibits are conceptualised in terms of their proximity to a putative 'ideal' exhibit as measured by their and the ideal's perceived characteristics; the more characteristics a real-world exhibit shares with the ideal, the more visitors it will attract. The theory seeks to provide an explanation as to why certain exhibits attract visitors and why others fail to do so.

A novel system of recording the behaviour of visitors to the Hall of Human Biology at the Natural History Museum using closed-circuit television and a real-time event recorder is described; and a study of inter-observer reliability is reported. A large prospective study of visitors to the Hall of Human Biology is reported in which the 'attracting power' of exhibits are defined and measured. In addition new statistics pertinent to the behaviour of individual visitors are defined and related to individuals' interest in the topics covered in the exhibition.

Further studies report how the characteristics of exhibits as perceived by visitors were elicited; and how a sample of real-world exhibits and the ideal were evaluated by visitors in terms of these characteristics to provide the necessary data for testing

the theory.

The implications of the theory for designers of exhibitions are discussed and further avenues for research are suggested.

	Page
CHAPTER 1	1
CHAPTER 2	15
CHAPTER 3	39
CHAPTER 4	50
CHAPTER 5	71
CHAPTER 6	88
CHAPTER 7	133
CHAPTER 8	165
APPENDIX I	184
APPENDIX II	195
APPENDIX III	199
BIBLIOGRAPHY	209

LIST OF TABLES	Page	LIST OF FIGURES	Page
Table 1	79	Figure 1	11
Table 2	81	Figure 2	31
Table 3	82	Figure 3	56
Table 4	84	Figure 4	57
Table 5	85	Figure 5	60
Table 6	98	Figure 6	65
Table 7	100	Figure 7	89
Table 8	105	Figure 8	106
Table 9	108	Figure 9	107
Table 10	111	Figure 10	109
Table 11	112	Figure 11	110
Table 12	115	Figure 12	116
Table 13	118	Figure 13	117
Table 14	121	Figure 14	124
Table 15	122	Figure 15	127
Table 16	126	Figure 16	152
Table 17	128	Figure 17	157
Table 18	136		

CONTENTS: continuedLIST OF TABLES

Table 19	List of Characteristics describing Exhibits in the Hall of Human Biology	140
Table 20	Randomised list of items used in Study II	147
Table 21	The percentage endorsement of items for ideal exhibit and five exhibits 'close' to the ideal and five distant to the ideal	151
Table 22	Relationship of items of Ideal Exhibit	153
Table 23	Reduced list of items from the original list of 53 items used for measuring the perceived characteristics of exhibits in the Hall of Human Biology	159

LIST OF FIGURES

Figure 18	Factorisation of exhibits across 23 items	158
-----------	---	-----

CHAPTER 1

INTRODUCTION

On May 24th 1977, the Hall of Human Biology opened at the Natural History Museum in South Kensington. The exhibition, occupying 1,200 square metres of the west wing on the ground floor of the Museum, is the largest to be opened this century; and since 1977 a number of other new permanent exhibitions have also been opened, namely Introducing Ecology (1978), Dinosaurs and their living relatives (1979), Man's Place in Evolution (1980), and most recently of all Origin of Species (1981). All these exhibitions are the fruits of a new exhibition scheme whose seed was sown early in the 1970's.

The Hall of Human Biology is of particular interest to psychologists because a large part of it included exhibits on psychology and members of both the British Psychological Society and the Experimental Psychology Society were consulted during its design and construction. Furthermore, for the first time in a museum in Britain, psychologists were employed to advise on the application of psychological principles thought to be relevant to the design of exhibitions, and to evaluate the visitors' responses to it. In fact, this thesis has arisen as a direct result of the research into visitors' behaviour and cognition in the Hall of Human Biology.

In this chapter I shall describe firstly the background to the project, followed by a description of the Human Biology exhibition and the approach to its design, and conclude with a few words on

the way in which the present research developed.

Background to the Exhibition Project

The Natural History collections of the British Museum were housed in Bloomsbury until 1882. The chief instigator and moving spirit behind the foundation of the new Museum in South Kensington, built between 1873 and 1881, was the great Victorian palaeoanatomist and advisor to Bishop Wilberforce, Sir Richard Owen (1804 - 1892). He was responsible for both the existence and policy of the Museum as it is known today (Miles and Tout, 1978). The policy may be stated briefly as having the twin aims of "diffusing and advancing knowledge about the natural world." To this end, the collections were initially arranged in three distinct series, each with a specific purpose. The first was an Introductory Series "by which the study of every group should commence"; the second was an Exhibited Systematic Series "in which the more important types are shown"; and the third a Study Systematic Series "required for enlarging the boundaries of knowledge." The galleries were laid out according to the plan of the Creator; all living fauna, the birds of the air and the fishes of the sea being placed to the left of the entrance way (the Hall of Human Biology has replaced the original fish collections); while to the right are the fossils buried deep, as it were, in the bowels of the earth, above them the minerals and, as in God's Plan, the plants above these on the third floor. Vestiges of this arrangement can be seen to this day. Owen quite explicitly it seems, acknowledged that natural laws and processes could not be demonstrated satisfactorily in a museum. At a time when the objects of biological enquiry were just that - tangible objects such as minerals, fossils and living animals - this restriction in what could profitably be displayed was barely a restriction at all.

However, over the years, scientific enquiry has turned increasingly to less tangible subject matter such as the study of genetics and behaviour. Accordingly, the restriction not to show laws and processes has resulted in the Museum being less and less the display of modern biology it is intended to be and once was.

Spasmodic modernisation has led to considerable change in the public galleries, and the study collections have been considerably enlarged by the fruit of numerous scientific expeditions.

However, by 1970 it became clear that change in the galleries was necessary and an Exhibition Scheme was begun. This scheme is described in detail by Miles and Tout (1978) and Miles and Alt (1979). For present purposes, it is important to note that the scheme is new in both the scope of the content and the manner of its presentation.

The scope of the new exhibitions covers all forms of life - including Man who was largely excluded from the subject matter of the original displays. However, the intention is to do more than just demonstrate the diversity of the living world, to include such topics as ecology, psychology and anthropology. Four main themes will be covered in the exhibitions as a whole: Man, Ecology, Evolution and Diversity and Life Processes and Behaviour. Although rumours have spread that the Museum intends to do away with many of its more famous exhibits, this is not so. However, such has been the outcry among academics about the approach and content of the new exhibitions, that in response the Museum felt it necessary to restate its policy in May 1981. Human Biology now represents the complete Man theme; and in future the Museum will concentrate on displaying its traditional subject matter - the diversity of nature. The other themes will now be subsumed in major exhibitions on diversity although it is still planned

to mount a separate exhibition on genetics at a later date.

The Human Biology Exhibition

The development of Human Biology owes a lot to the approach fostered by North American "science centres" (Kimche, 1978). The Exploratorium in San Francisco built by Frank Oppenheimer (Oppenheimer 1968) and the Ontario Science Centre in Toronto (Wilson 1976) have shown convincingly that it is possible to display the principles as well as the paraphernalia of science. In line with these institutions, in the Hall of Human Biology, the old, formal rows of dusty cases have been banished; there are many large models which visitors can touch and explore, but few glass cases. There are slide shows about prenatal development, about aspects of neuro-endocrinology, about language acquisition and several other topics. Games display the control of behaviour by the brain, learning and the problems of logical reasoning. Such untraditional exhibits are not to everyone's taste but they are an honest attempt to show that what matters is, for example, not just what a brain looks like, but what it does; or more generally, an effort to go beyond the mere appearances of things scientific.

Much thought was given to the subject matter to be included in Human Biology for clearly there is much more to be said on the topic than can be covered in the space available. The exhibition includes the minimum number of concepts consistent with the overall story which traces the development of a fertilised ovum to a fully mature human adult. All the concepts were organised into a learning hierarchy (Gagné, 1970) that made no unrealistic assumptions about the visitors' vocabulary and previous knowledge. Simple ideas are dealt with first; these

form the building blocks for more complex concepts. Technical language is avoided except where absolutely necessary; and the sheer amount of text is kept to a minimum.

The exhibition is divided into physically independent sections, each one dealing with a set of related concepts that are part of the learning hierarchy. The ten major sections shown in Figure 1 are as follows:

- A. Living Cells
- B. Growing
- C. Movement
- D. Controlling your actions
- E. Homeostasis - Your life in the balance
- F. Hormones
- G. Hormones and Nerves
- H. Experience of a lifetime - Learning and Memory
- I. Perception - Understanding the world about us
- J. How do we come to understand the world about us.

In addition, further material relating specifically to particular sections is presented in separate side assemblies - "More about Chromosomes" (BE), "More about controlling your actions" (DE), "More about Sex Hormones" (FE), and "More about Memory" (HE). These areas are equivalent to the "enrichment" booklets of some modern school courses in the United Kingdom, and they are designed to provide the visitor with a deeper understanding of certain of the topics covered in the major sections.

The organisation and presentation of the subject matter of the exhibition has drawn heavily on a number of principles developed in the field of educational technology. These have been summarised by Miles and Alt (1979) as follows:-

Image redacted due to third party rights or other legal issues



1. the educational purpose is spelt out in the form of unambiguous objectives. These provide the basis for the acquisition of objective knowledge about Museum communications:
2. the subject is carefully arranged so as to present the underlying ideas in a logical sequence;
3. the interaction between the learner and the learning material is carefully controlled by:
 - (a) arranging the material in digestible steps,
 - (b) making provision for different capacities to learn and various levels of knowledge and interest, e.g. by including "enrichment" material,
 - (c) making provision for participation or active responding, though not necessarily through overt responses,
 - (d) making arrangements for immediate feedback, i.e. for the learner to discover at once whether or not his response has provided an acceptable answer to the question or solution to the problem,
4. the educational medium (audiovisual, diagram, three-dimensional object and so on) is chosen because of its appropriateness both to the message to be conveyed and to the intended recipient;
5. the educational materials are tested to see whether they produce the results specified in the objectives. These materials are the physical entities which correspond to the testable hypotheses of scientific research. If a test is properly designed, the results

will show whether or not the materials are suitable for their purpose. More importantly, it will provide information of more general value for later use.

Human Biology set out to involve visitors actively, making a striking break with the traditional natural history museum. The exhibition contains a considerable number of "interactive" exhibits. Such exhibits are not simply those where the visitor is expected to press a proverbial button; they are exhibits with which the visitor is encouraged to participate - cognitively. In short, the exhibition aims at presenting scientific information in exciting and stimulating exhibits, avoiding the traditional unsympathetic juxtaposition of visitors and facts (Wittlin, 1971). A passive role is no longer imposed on visitors; they are encouraged to seek new information by exercising their own judgement in answering questions, by operating various interactive devices, and by playing games designed to model various scientific phenomena.

Research among visitors to the Hall of Human Biology

Perhaps the most interesting aspect of the Human Biology exhibition is that from the outset it has been thought of as a pilot exhibition in which exhibits and their effects on visitors were to be assessed. In fact, it is confidently expected that within fifteen years from its opening date, half of all exhibits in Human Biology will have been replaced by "better" exhibits. For example, as a result of criticism (Gregory 1978), the "Perception" section was completely redesigned and a new version opened early in 1982, and the "Homeostasis" area was redesigned in 1980 as a consequence of an unreported observational study of visitors to the area.

The majority of this thesis is devoted to reporting the major findings of a number of studies carried out during the period July 1977 - April 1981 among casual visitors above the age of sixteen years to the exhibition.

The prospect of carrying out research among visitors to the Natural History Museum is interesting to psychologists for a number of reasons. Policy at the Museum precludes the possibility of adopting an experimental approach. The only interventions a researcher is permitted to make in the natural course of an individual's visit to the exhibition is to question him at the beginning, during or at the end of his visit; no experimental manipulations of the exhibits looked at or the routes taken through the exhibition are possible. Therefore, by necessity, the research reported is essentially "naturalistic" and not experimental; although in certain instances experimental procedures were possible.

Although there is a large number of reported studies of museum visitors, the overwhelming majority of them have been carried out without reference to current psychological theories; they resemble "market research" in the sense that the data collected are atheoretical. It is a considerable challenge to incorporate museum research into mainstream theoretical psychology.

Finally, in reporting the research, a narrative structure has been used. This is not meant to imply that the ideas developed in quite the same orderly fashion in which they have been presented. As Harré (1972) has pointed out, most research is a leap-frog process of fact accumulation and theory generation. This research is no exception but in the interests of the reader, studies which led nowhere or up blind alleys have been omitted as they are irrelevant to the main story.

CHAPTER 2

LITERATURE REVIEW

There has now been an enormous amount of published research on museum visitors. For convenience, the methods used can be described under three broad headings -

- (1) large-scale sample surveys,
- (2) studies of behaviour, and
- (3) educational evaluation of museum exhibits
using paper and pencil tests of knowledge
(Alt and Morris, 1979) -

although, in practice, a combination of methods has sometimes been employed. In this chapter research in each of the three categories will be described, followed by a critical discussion.

Surveys

It has been estimated that approximately two-thirds of all published research fall into this category (Loomis, 1973 a), and the majority are carried out to provide museum planners with detailed demographic information about their visitors and their reasons for visiting. A bibliography compiled by Elliot and Loomis (1975) gives numerous references to visitor surveys, the majority of which tell us that visitors come to museums out of general interest, are such and such an age, social class and distance from the museum; that they heard about the place from

friends, radio or television in varying proportions; and that they 'enjoyed' their visit, and so forth. At the Natural History Museum a series of annual visitor surveys was initiated to monitor the success of the exhibition scheme. Some measure of the success of new exhibitions can be gauged by the number of people who mention them as a reason for visiting the museum, or as exhibitions they particularly enjoyed, or can remember most about. This sort of information can be best collected by repeated study over time. The first five annual surveys have been reported (Alt, 1980) and a copy of this paper is included in Appendix 1, primarily as it describes in detail a systematic sampling method that has been used throughout the present research project for selecting subjects. However, in general, surveys are of little interest to psychology as they are atheoretical in their approach.

Behavioural observation

Systematic observations of the behaviour of museum visitors apparently began during the 1920's in American art galleries and these studies have been reported by Robinson (1928, 1930, 1931 a, 1931b, 1933 a, 1933 b, 1933 c). It was found feasible for observers armed with stopwatches, pencils and paper to follow individual visitors inconspicuously as they ambled through a museum, recording the total time spent in the museum, the areas visited and the length of time spent looking at particular objects. A number of general patterns of behaviour emerged. For example, during the course of a visit and after a brief 'warming up' period the persons observed displayed a tendency to stop before a progressively smaller percentage of the pictures encountered and to make progressively shorter stops; the more pictures displayed on a gallery wall, the smaller the average time spent

looking at each individual picture; and, in large museums, the likelihood that a visitor will observe any given picture is less than in smaller museums. Robinson's work in art galleries was followed-up in the 1930's by his student Arthur Melton who later became better known in mainstream psychology through his work on human memory (see Melton 1933 a, 1933 b, 1935, 1936, 1972; and Melton, Goldberg and Mason, 1936) and similar work was conducted in the Peabody natural history museum by Porter (1938).

Consolidating and building upon Robinson's work, Melton discovered the pronounced 'right-turning tendency' prevalent it seems from later studies in all American museums. He also described a so-called 'exit-gradient', whereby time allotted to a painting seems to be directly related to its physical proximity to the exit door. He found a close relationship between, on the other hand, attractiveness (defined as the probability of stopping at a painting) and density of paintings on the wall (which seems to obey a kind of Weber-Fechner law); and on the other hand, the independence of 'holding-power' (defined as the time spent at a painting given that the visitor has stopped moving) and display density. Melton also described results showing that 'museum fatigue' affects attractiveness while leaving 'holding power' unaffected, and that such fatigue is almost certainly a kind of object fatigue termed 'ennui' by Melton - rather than sheer physical fatigue, because the steady decline in attractiveness from the time of entry sets in immediately and can be quite pronounced, even in visits as short as five minutes.

There was something of a hiatus in visitor research of any kind until the 1960's when a new interest burgeoned; and this was manifested by the emergence of a large number of published papers. During this period a number of studies were reported that were

firmly within the paradigm established by Robinson (for example, Abler 1968; MacBrier 1964; Munyer 1969; Parsons 1968; Parsons and Loomis 1973; Weiss and Boutourline 1963) but it was not until the work of Winkel and Sassanof (1970) and Lakota (1975) that any new developments emerged. Winkel and Sassanof built a simulation-system which adequately modelled the behaviour of museum visitors in a real gallery while subjects sat watching a triple slide-screen depicting the view to the left, straight ahead, and to the right, respectively. Through instructions to the slide-machine, subjects could 'move' through the gallery. Perhaps somewhat surprisingly, the routes taken mapped onto the observed routes in the real gallery well. Lakota has described a very detailed behavioural analysis in the natural history section of the Smithsonian Museum in Washington. Although the procedures he employed are extremely complex to understand, he appears to have built a quite sophisticated regression model of attractiveness, holding power and effectiveness (a kind of composite of holding-power and amount learned at the display.) Other equations involve different factors, one of which indicated that the Beta coefficient for modern display methods has a high negative value on holding-power!

Educational Evaluation

Historically, the educational evaluation of museum exhibits is rooted in the work on curriculum development in America which gained momentum during the 1960's. During this time, which saw the rise of programmed learning, a considerable amount of attention was given by educators to 'instructional objectives' which became widely known as behavioural objectives. These described explicitly what it was hoped learners could do after

instruction that they could not do before, and they were contrasted with the loose statements of objectives that had been previously used in education.

Shettel (1968) suggested that instructional objectives of the sort outlined for use in mainstream education by Bloom et al (1956), Krathwohl et al (1964) and Mager (1962) should collectively form the basis for exhibition design and evaluation. His prescription is that 'exhibition objectives' should be fixed in advance and design should then proceed accordingly. Data are then to be collected from samples of visitors to estimate the extent to which these objectives have been met and exhibitions redesigned in the light of the findings. In the field of curriculum evaluation, Scriven (1967) drew a distinction between formative and summative evaluation. These terms refer to when rather than how evaluation is carried out and both are relevant to Shettel's idea of the value of exhibition objectives. Scriven (1976), adopting the curriculum evaluation terminology, has recommended that exhibits should be developed by a process of formative evaluation and he has provided a prescriptive process for doing so; and a number of exhibits developed along these lines have been reported in the literature (for example, Eason and Linn 1976; and Scriven and Brown 1978).

The contribution of Shettel

The most ambitious project strictly along the lines laid down by educational evaluators has been reported by Shettel et al (1968) in which the results of an eighteen-month investigation of "The Vision of Man" exhibit, sponsored by the U. S. Office of Education are described. The exhibit covered non-military scientific and technical programmes conducted by the United States

federal government, sometimes in conjunction with private industry; it also outlined the accomplishments of these programmes. Although designed to appeal to a wide audience, it was aimed at secondary-school children in particular, who might be influenced in choosing a science career with the federal government. The exhibit employed a wide variety of different design techniques using modern exhibition technology. As such Shettel considered it representative of a modern, sophisticated state-of-the-art display, and thus an appropriate medium in which to test the potency of a variety of effectiveness measures.

In the 1968 report it was emphasised that the purpose of the research programme was not to evaluate "The Vision of Man", per se, but to evaluate the methods used to perform the evaluation.

The exhibit was conceptualised as consisting of three sets of variables:

- (1) exhibit design variables,
- (2) exhibit viewer variables, and
- (3) exhibit effectiveness variables.

The first were considered as independent or experimental variables, the second as control variables, and the third as dependent variables. The investigation set out to identify the important variables within each set of variables, their relationship with each other, and the best ways to measure them.

The exhibit design variables that were evaluated were physical characteristics of the exhibit itself, such as the readability level of labels, location of various parts, etc., the exhibit viewer variables included audience characteristics such as age, educational level, etc., the effectiveness variables included

observations of viewers in order to arrive at measures of attracting and holding power, and various tests designed to measure knowledge, attitude, interests, etc.

The project involved two major stages - the exhibit-testing stage and the mock-up-testing stage. The exhibit-testing stage was used to devise, use, and revise the measures of empirical effectiveness as well as to collect data regarding the relationships between the different types of variables. During this stage no experimental variations were made in the exhibit design variables but they were made in some of the exhibit viewer variables. Specifically the variables 'manipulated' were age, education, socioeconomic status, sex, amount of science background, the extent to which exhibit viewing was voluntary, and viewing time.

The exhibit mock-up stage was used to manipulate several of the exhibit design variables - the simplification of text, the amount of visual illustration, and the use of audio devices to replace text portions of the exhibit.

Each of the two testing stages comprised two phases, separated in time. The primary objectives of exhibit-testing phase 1 were developing the initial versions of the effectiveness measures, collecting data on the usefulness of these measures and revising the measures. During phase 2, data were collected on the revised measures of effectiveness and these data were used to generate hypotheses about the relationship between the three sets of variables referred to earlier. The first phase of the mock-up testing stage was concerned with validating the mock-up exhibit and with collecting data on experimental variations for several design variables. Phase 2 was devoted to collecting data on

experimental variations in several exhibit design variables such as the amount and readability of text, the amount of visual illustration, and the use of audiovisual textual communication.

Shettel's first task was to translate the original objectives of "The Vision of Man" into measurable objectives prior to any test development - they were not originally stated in terms that readily lent themselves to measurement. The form these translations took was almost exclusively knowledge-oriented because, according to Shettel, "... this is the only feasible way of measuring such objectives short of long-term follow-up studies of viewer behaviour (and even this approach would be subject to gross contamination)."

A post hoc analysis of the exhibit content was then carried out. This entailed numbering and locating all the parts of the exhibit, performing a readability analysis of all text and labels, noting type size, colour, etc. At the same time colour photographs were taken of all static models and motion pictures were taken of all moving parts. The information collected at this stage was used to construct the simulation of the exhibit for the mock-up-testing stage as well as to provide measures of certain independent or experimental variables.

The next stage was devoted to developing background, interest, and attitude measures. The background questionnaire contained questions about age, educational background, socioeconomic class, etc., as well as subjective measures of the amount of time spent at the exhibit, whether a reason for the museum visit was to view the exhibit, awareness of the exhibit in the media, and intentions with respect to future career. The interest measure was based on the assumption that if people said they were interested in a

particular area or topic, they would be able to remember something about it. A small questionnaire was administered to measure reported interest. The attitude measurements were also collected; the items in the questionnaire were based wholly on the original objectives. Each item was devised to reflect viewer opinions and attitudes towards one objective.

Developing the knowledge measures received the most attention. Four individual questionnaires were designed to measure the amount of information visitors retained after viewing the exhibit. These measures were based on the assumption that there are various levels at which learning can be measured. Two recall methods were used to identify a high level of learning on open-ended questionnaires. Two recognition tests were used to measure lower levels of learning on multiple-choice questionnaires. One of the open-ended questionnaires asked about open-ended concepts and the other about open-ended knowledge; the recognition tests were split in a similar manner. As the names imply, the concept questionnaires were designed to measure the visitor's ability to recall and recognise scientific concepts or principles, and the knowledge tests were designed to measure the specific learned information that could be recalled and recognised.

As mentioned earlier, the exhibit effectiveness variables included observing visitors to arrive at measures of attractiveness and holding power. A number of approaches were adopted in "The Vision of Man" study to answer the question: "Is there a correlation between the number of static and dynamic modes and the attracting power of the exhibit." Visitors were followed (unobtrusively), and a videotape recording system was also employed to monitor behaviour at individual displays. These methods were felt to be best in meeting the requirements of unobtrusive observation.

The survey design incorporated four groups of subjects,

- (1) a max. group
- (2) a min. group
- (3) a control group, and
- (4) a casual visitor group.

The first three groups comprised volunteer subjects who were paid for participating in the project, and the fourth group was made up of museum visitors who came to the exhibit of their own volition. The max. group subjects were given instructions to spend as much time as they wished in the exhibit with instructions to learn as much as possible because they would be tested at the end of their visit. The underlying aim in instructing this group was to discover how much could be learned from an exhibit given maximum amount of time and motivation. Conversely, the min. group, designed to establish a minimum index of learning, spent only half an hour in the exhibit; they were also given instructions to learn as much as possible. The control group subjects, comprising people who had not seen the exhibit, were also tested to establish baseline scores on all the test measures. The members of these three groups were matched on age, sex, education, and science/non-science background, and different subgroups within these categories were represented. The casual visitor group subjects were invited to participate in the testing procedures; some were interviewed before they commenced their visit and others were interviewed at the end of their visit. All four groups were given the batteries of tests described earlier but, obviously, only the casual visitor group was used to establish the exhibit effectiveness variables.

The mock-up subjects were high-school students who viewed thirty-one individual panels, designed to represent all static sub-areas

of the actual exhibit. On the panels were mounted colour photographs of the exhibit sub-areas with associated text. The introductory panels of the mock-up were hung together in one room; all other mock-up panels were hung consecutively along the main corridor walls. Subjects were randomly assigned to max., min., and control groups similar to those previously described.

Not surprisingly, the results of the investigations showed that the amount of viewing time and motivation¹ influenced the amount of knowledge gained from the exhibit, as reflected by subjects' scores on the various test items. Thus, max. group subjects always achieved higher test scores than min. group subjects, who, in turn, always scored higher on the test items than did control group subjects. The actual museum visitors, who came to the exhibit on their own accord, scored lower on the test items than did control group subjects! In Shettel's opinion, these results show that "The casual visitor group, as a whole learned very little from the exhibit as measured by the tests used." Generally, college students scored higher than either high-school children or adults; subjects with a science background scored consistently higher than subjects without one.

Results on the interest measures showed that viewing the exhibit did influence the level of interest in the subject matter, although no consistent patterns emerged. Results from the attitude measures were largely inconclusive; all the experimental groups obtained scores relatively similar to the casual visitor group.

¹ Motivation was never explicitly defined or varied. It is presumed to refer to the experimental subjects' assumed desire to do well on the tests they knew they would take at the end of their visit.

Unobtrusive observations of casual viewers showed that the average amount of time spent at an individual display was twenty seconds, and the total time spent in the whole exhibit area was fourteen minutes. A technique for measuring attracting power involved the use of a multiple-regression equation based on:

- (1) the number of subjects actually stopping at each display,
- (2) the number of static parts, and
- (3) the number of moving parts.

This equation resulted in an attractiveness rating for each display. The results indicated that the number of static and moving parts in a display accounted for twenty-seven per cent of the variance in viewer stopping behaviour; however, the number of moving parts in a display was a much better predictor of viewer stopping behaviour than the number of static ones.

Analysis of the videotape recordings of casual visitor behaviour at individual display permitted the calculation of relative attracting power among them. In general, dynamic parts and their associated text were looked at by casual viewers much more frequently than static pictures and their text.

Attempts at relating the reading difficulty of text to the amount of knowledge gained were inconclusive. Thus, the results did not indicate that simple text facilitate any more learning than did more difficult text.

The results of the exhibit mock-up testing (phase 1) showed that the mock-up subjects performed in a comparable manner to the exhibit subjects on the various tests. That is, the results obtained by max. subjects were similar to max. mock-up subjects,

min. exhibit subjects were similar to min. mock-up subjects, and so on. These results were taken to indicate the validity of the simulation technique as a means of judging actual exhibits. The experimental variation in the mock-up design (phase 2) indicated a significant difference favouring the group reading the text but no difference between the other experimental variations.

In the discussion section of Shettel's report, much space is devoted to reiterating previously stated views, such as the need for unambiguously stated exhibit objectives, the appropriateness of various evaluation tests, the use of mock-up exhibits in developing didactic exhibits, etc. The penultimate section presents what is grandly termed a "three-factor theory of exhibit effectiveness," suggesting that an exhibit must fulfill three purposes in order to be educationally effective:

- (1) it must initially attract viewers to the exhibit,
- (2) it must maintain their attraction throughout the exhibit, and
- (3) it must maximise the amount of relevant learning or influence what is achieved on the part of the viewer.

Finally, Shettel suggests that results of his study indicate several hypotheses regarding the effective organisation of didactic exhibits that could be explored in future projects.

Shettel's work has been described in detail because it represents the typical approach to evaluation and exhibition design currently extant among educationalists. However, Screven (1975) has gone even further than Shettel by applying the technique of programmed

learning to museum education.' Visitors were given booklets of questions to look at as they went through an exhibition, or a cassette tape recorder, together with electronic response cards with which to answer the questions. Visitors successfully finishing the course were then rewarded with a badge developing their status as 'museum experts'.

Critique

Criticism of museum research may be raised on three broad fronts;

- (i) methodological,
- (ii) philosophical, and
- (iii) the utility value, or usefulness, of the research to designers of exhibitions.

(i) Methodological Critique

In a critical discussion of visitor surveys, Loomis (1973 b) complained that they varied from very bad (e.g. poor sampling, badly worded questions, etc.) to excellent (e.g. good survey practice in every sense, a "model" of visitor attendance guiding question design, etc.). In surveying the literature on observational research and educational evaluation with particular reference to the studies in the bibliography compiled by Elliot and Loomis (1975), there is no estimation of the reliability of the various measurements taken and an inadequate consideration of their validity irrespective of whether the research is observational or pencil and paper tests.

The reliability of the methods used to observe visitors' behaviour

will be discussed in detail in Chapter 5.

Robinson and subsequent researches focusing on the behaviour of museum visitors have argued that attractiveness and holding power are valid measures of exhibit effectiveness. An effective exhibit must first of all attract visitors and then hold their attention for a sufficient time to impart information. Miles and Tout (1979) have suggested that holding power can be extended to include an index of required viewing time. They recommend an index they call 'holding power ratio', i.e. the actual viewing time divided by the required minimum viewing time. Thus a score of less than one would mean that visitors were spending less time than is necessary to receive the exhibit's message. Educational evaluators (e.g. Lakota 1975; Lakota and Kanter 1976; Screven 1976; Shettel 1968 a, 1968 b, 1973), while accepting that high attractiveness and high holding power are necessary for an exhibit to be effective, claim that they are not sufficient by themselves. For an exhibit to be truly effective it must communicate the message/s intended by the designer. Hence their interest in measuring the knowledge and understanding imparted by exhibits. Educational effectiveness is measured by pencil and paper tests and judged by the extent to which certain external criteria (e.g. a pre-stated percentage of visitors achieving a preset percentage score on the tests) are met. In these circumstances it is usual for validity to be considered in terms of criterion-related validity. Presumably, the higher the correlation between attractiveness, holding power and the scores on the tests, the better the validity of the tests. Nowhere in the studies reported in the literature has validity been considered in this way and in view of the fact all show that casual visitors learn nothing, as measured on the various tests, it follows that the validity of the tests must be called into question.

It is self-evident that attractiveness is a valid measure of effectiveness; if a visitor does not look at an exhibit the exhibit cannot be effective. The issue with holding power is only a little more complicated. As Miles and Tout (1979) have pointed out, it seems reasonable to suppose that holding power is related to the characteristics of the exhibit such as length of text, and it was for this reason that they formulated 'holding power ratio' as an alternative but related measure.

There is another way, however, in which the validity of attractiveness and holding power may be considered and this is in terms of their validity as theoretical constructs in a theory of exhibit effectiveness; this is known as construct validity. As an approach to the question of their validity, it will be taken up in the following section.

(ii) Philosophical Critique

The research workers whose studies have been described earlier in this chapter all set out to evaluate the effectiveness of museum exhibits in a self-consciously scientific manner; that is, they claimed to have used the methods 'characteristic' of scientific enquiry. Essentially, the reported studies all follow the same set of procedures. Firstly, they begin by observing and collecting certain 'facts' about exhibits and visitors. For example, Robinson and Melton were interested only in whether visitors stopped at an exhibit or not, and if so, for how long. As described earlier, Shettel collected a large number of facts about exhibits and visitors. After the facts have been collected they are analysed and classified; and then, by a process of induction, generalisations are made. In certain cases, (e.g. Melton's work) the

generalisations are tested by further experimentation or observation. In fact, the studies follow the prescriptions of correct method given in many social science textbooks. For example, Figure 2 is taken from Wallace (1969). Its components are four stages linked by four methods in a clockwise process.

Image redacted due to third party rights or other legal issues



Simple though the diagram is, it represents the approach to visitor research very accurately. In the philosophy of science this characterisation of the scientific process is known as Positivism.

Epistemologically, the positivist conception of the nature of scientific investigation will not stand serious criticism. Firstly, the notion that scientific knowledge about effective exhibits is advanced by the accumulation of facts independently of specific

prior hypotheses about their relationships is simply untenable.

By way of analogy, the work of Gregory (1966) has shown that visual perception is not determined simply by the stimulus pattern on the retina but that it involves actively searching for the best interpretation of the sensory data. This hypothesis formation - to use Gregory's terminology - involves not only the sensory information but also knowledge and experience. Gregory (1980) has drawn a deep parallel between the nature of perception and the nature of hypotheses in science. Hypotheses in science give direction to scientific investigation just as prior knowledge and experience direct the perceptual process. The way to ensure that the meaning of a fact advances our knowledge is to place it in the context of an antecedent hypothesis. The alternative is often to create confusion by collecting a jumble of unrelated bits of information with unfathomable meanings. Shettel's work is a good example; he collected numerous 'facts' without any clear-cut hypotheses about their relationships; and very little by way of substantive knowledge resulted.

The problems with induction as a form of reasoning are very familiar (see for example, Harré 1972), and are glaringly illustrated in visitor research by the conspicuous absence in all so-called theories of exhibit effectiveness of any explanation of what makes for an effective exhibit. A theory of exhibit effectiveness should at least explain how exhibits differ in terms of their effectiveness and should make an attempt to explain how the factors that affect effectiveness do so. The concept of construct validity mentioned in the previous section has a direct bearing on this issue. For example, what psychological or other property or properties can explain the variability in attractiveness and

holding power between exhibits? No satisfactory explanation, or answer to this question, can be given in the absence of a theory of exhibit effectiveness to underpin the measurements that are supposed to indicate effectiveness. Of course, the operationist view which accompanies the prescription epitomised in Figure 2 holds that measurement can proceed independently of theory; and, if we follow the methods laid out in Figure 2, it should precede theory.

But, as Hull (1968) argued, and has been suggested here, it is necessary that theory precedes measurement for satisfactory explanations of empirically derived results to be forthcoming.

In environmental psychology Proshansky et al (1970) have characterised examples of two different approaches as the 'physical' approach and the 'phenomenological'. The physical approach, it is claimed, is the objective one whereby attempts are made in any given environment to relate man's behaviour to discrete and quantifiable stimuli present in that environment. Deutsch and Krauss (1965) have described the same general approach in social psychology as follows:

"Just as the animal psychologist who is studying the behaviour of a rat in learning a maze must know the relevant physical properties of the maze to understand or predict the behaviour of the rat in it, so too the social psychologist must be able to characterise the relevant features of the social environment in order to understand or predict human interaction".

In other words, the 'physical' approach adheres to the principles of Watsonian behaviourism.

The phenomenological approach; on the other hand, has sought to account for man's behavioural relationships with the environment in terms of the interpretations and cognitive constructions by which he or she organises his or her experiences of that environment.

Behavioural studies of museum visitors fall clearly within the physical approach; and those studies concerned with educational evaluation, while not 'physical' in Proshansky's sense, are still mechanistic in conception. For example, Shettel has argued that educational exhibits satisfy the 'need' to learn. This is an organismic view of man in which cognitive functioning is conceptualised in terms of biological needs similar to the needs for food and water, for example. Presumably, visitors are thought of as arriving at exhibitions in a state of ignorance and upon being confronted by didactic exhibits their 'need' to learn operates as a drive and is reduced by the consumption of the knowledge on display.

All theories of exhibit effectiveness attribute to the visitor an essentially passive nature and the mark of them is to treat visitor action as a natural and determined phenomenon. Apart from any random factors (giving rise to probabilistic laws) the visitor accordingly behaves predictably in given conditions and he can be manipulated by engineering appropriate conditions (that is, effective exhibits). Thus, the individual visitor is not causa sui in the explanation of his own behaviour. Exhibit learning, the routes taken through exhibitions, exhibit interest, and so on, are dependent variables, functions of 'drive states' or exhibit design.

Compton (1935) pointed out, if completely determined laws are

applied to man's actions, he is himself an automaton; and Gauld and Shotter (1979) have suggested that if a deterministic view is upheld, then the possibility of a scientific or experimental study of man's behaviour is logically questionable.

"If a scientist is himself no more than an arena within which inevitable causes produce their inevitable effects, how could he have discovered the causal laws which his own behaviour is supposed to instantiate? A scientist practising his craft must assume himself to be a free agent, a being capable of manipulating his corner of the universe. The effect of his manipulations must be to destroy any chance regularities among phenomena leaving the real, i.e. causal regularities to show themselves. But if the scientist's intervention is itself causally determined the possibility must remain that his intervention did not destroy a fortuitous regularity but was the upshot of an antecedent state of affairs which also caused the disruption of the regularity."

Miles and Tout (1979) have argued that the issue of whether or not man's action can be regarded in a deterministic manner is surely of theoretical and not practical interest. Their view misses the important point that practical issues are underpinned by theoretical standpoints. Accepting that museum visitors actively select the exhibits they look at, rather than being jerked into life by them, can have a fundamental impact on how one judges the existing indices of exhibit effectiveness. Holding power has become part of what Lakatos (1974) would have called the 'hard-core' of any theory of exhibit effectiveness and it is a construct which presupposes a passive conception of the museum visitor. Effective exhibits are thought of as those

that hold visitors' attention long enough to get across their messages. It is proposed that if an exhibit is designed effectively, visitors will remain at it for as long as the designer intended. If visitors do not stay the required time the exhibit has been by definition, badly designed. No mention is made of whether visitors are willing to invest the necessary time to receive the message at a given exhibit; it is assumed that their passive nature can be manipulated by effective exhibits which generate the required holding power. Apparently there is no upper limit to the amount of time visitors can be made to remain if the exhibit is effective. And yet, all the research that has been published to date has shown that visitors, on average, spend less than a minute at individual displays; and this finding is upheld in different exhibitions and different museums. If one upholds an active conception of the nature of the museum visitor, one would argue in the face of such overwhelming evidence that perhaps visitors are not prepared to spend longer than thirty or so seconds of their time at an individual display; and that seen from the visitor's point of view rather than the designer's an effective exhibit is necessarily one that gets its message over very quickly. Thus, holding power when considered within a phenomenological framework becomes a suspect empirical concept.

Utility Value

The utility value of visitor research can be measured, in theory, by the extent to which the findings can be applied to the solution of practical problems in designing exhibits. The application of research findings to improve the effectiveness of museum exhibits has been termed museum technology by Melton (1935). More recently it has been resurrected by Oppenheimer (1968), Miles and Alt (1978) and Miles and Tout (1978, 1979). However, in the latter case it has been suggested that this technology should

include any objective knowledge from appropriate fields such as educational research and educational technology.

The information from large-scale sample surveys is similar to many market research investigations where data are collected about the characteristics of consumers of particular products, their attitudes towards them, their preferences, etc. with a view to modifying the products to match the needs of the consumers; and basic market research of this nature would be extremely useful to museum planners if they were prepared to act upon the results. In the absence of any such intent, visitor surveys have no utility value.

The results from behavioural studies have led Lakota and Kanter (1976) to recommend simple precepts concerning exhibition architecture which they claim have a profound effect on visitor behaviour. These recommendations have also been summarised by Miles and Tout (1979), as follows:

1. the exhibition space should be partitioned into smaller areas with chambers or alcoves large enough for groups of about ten to observe some aspect of the exhibition,
2. the organisational structure of the exhibition should be clear to visitors, and
3. island display units should be avoided or sited off the main routes through the gallery.

The results from educational evaluations have not by themselves given rise to any guidelines concerning the design of effective exhibits. Although a number of authors have suggested that principles developed in connection with curriculum development and

educational technology can be successfully incorporated in a museum technology (see Miles and Tout 1979 for a summary of the relevant principles that have been proposed) there is little or no empirical evidence to suggest that casual visitors learn more from exhibits incorporating these principles in their design from those that do not. It is taken as axiomatic, as part of the prevailing ideology, that these principles are applicable in the museum setting. By adopting the curriculum development model uncritically, there is an implicit assumption that visitors will want to learn from exhibits rather as they might from textbooks; and that visitors to exhibits will remain for a predetermined time rather like pupils in a lesson at school.

The major conclusions to be drawn from this chapter are that the studies reviewed have been carried out within a tradition of naive empiricism. 'Facts' about exhibits and visitors have been collected in the hope that patterns may be discerned within the data that arise from their categorisation and analysis; it is hoped that these patterns will in turn give rise to theories of exhibit effectiveness.

It has also been assumed that visitors have an essentially passive nature which can be directed and controlled by effective exhibits. Furthermore, related to the idea of a manipulable passive visitor, the underlying metaphysic or ideology in designing educational exhibits is that the principles developed within the disciplines of education and educational technology are directly applicable to the design of educational exhibits.

The present thesis rejects all these propositions in the process of developing and testing a theory of what makes for an effective exhibit that acknowledges the visitor as an active agent who consciously decides which displays he or she will look at.

CHAPTER 3

A THEORY OF EXHIBIT EFFECTIVENESS

Background

The natural history of a visit to an exhibition may be thought of as a series of interactions between the visitor and the individual displays he encounters during the course of his visit. A visitor's time is clearly limited and, therefore, it is unreasonable to suppose that he or she will be able to look at everything on display in the museum. An analogy may be drawn here between biological competition among individuals for any limited environmental resource and the competition between individual displays in their call on a visitor's attention. Successful displays are those that compete and attract the visitor; and unsuccessful ones are those that fail to do so. However, to resolve the tautology inherent in the analogy, a theory of exhibit effectiveness must offer an explanation as to the reasons why different displays are more or less attractive as well as making predictions about successful and unsuccessful displays.

The philosophical standpoint underlying the theory is that a visitor to an exhibition acts in accordance with his or her perceptions of the exhibition. In other words, from the visitor's point of view an exhibition is not simply a collection of 'objective' physical entities but a perceived environment, constructed by him. Piaget (1972) making a similar point stated:

"On the one hand, knowledge arises neither from a self-

conscious subject, nor from objects already constituted (from point of view of subject) which would impress themselves on him; it arises from interactions which take place midway between the two and thus involve both at the same time."

This standpoint contrasts markedly with those of the earlier researchers, whose work was reported in the previous chapter.

In Part I of his 1935 report, Melton included a section entitled The intrinsic attractiveness of the Art Object as a Determinant of Interest. In it, he claimed that the attitude of those concerned with the display of art objects could be epitomised in the assertion that "a painting will find its own level when it is displayed in a gallery along with other paintings." This, he asserted, is tantamount to saying that a painting will attract an amount of attention that is appropriate to its intrinsic worth. He suggested that the assertion was at best a pious hope and furthermore was based on assumptions of dubious validity. He contended that the attractiveness of a painting had nothing to do with its "aesthetic quality" as judged by experts. Paintings judged by experts to be of outstanding quality failed miserably in competition with other paintings. He went on to say that while he accepted that factors such as the subject matter, the size and the colouring of a painting appeared to be important in determining its attractiveness, they were not of primary importance. According to Melton, location and mode of display were of far greater importance. Indeed, Melton's work was concerned solely with these latter factors and, in particular, how they affected attractiveness and holding power.

In line with the arguments presented in the opening two paragraphs of this chapter, Melton's views must be rejected. Displays are

considered to be in competition with each other and the result of this competition is that each display will find its own level vis a vis the others. However, an acceptance of this does not amount to the same thing as claiming that a display will attract attention in proportion to its intrinsic worth or aesthetic qualities; and in suggesting that it does Melton was putting up an Aunt Sally. Characteristics or attributes of displays (what Melton called factors) do not inhere in the displays, as it were, but they are perceived by visitors interacting with the displays. Factors such as mode of display and location are only important in so far as they are perceived by visitors to be characteristics of a display; a display will attract attention in proportion to its perceived worth not its intrinsic worth. Of course, Melton is correct in stating that the attractiveness of a display does not necessarily have anything to do with the judgements of experts but in context this is a red herring.

Similar objections can be raised concerning the work of Lakota and Shettel. Unlike Melton, Lakota made an attempt to relate attractiveness and holding power to the characteristics or attributes of displays. Unfortunately, it is the most depressing section in his report. Noting the thousands of words written on the subject - "most of it inconclusive" - Lakota adopted a very down to earth pose. As an objective (my emphasis) scale of exhibit design, he suggested that the year in which an exhibition was designed is at least one unambiguous variable that can be inserted into the regression equations! Galleries designed in 1954 received a scale value of 1 while more recent galleries received higher values culminating in 5 for galleries designed in 1964. He also obtained ratings from three "museum professionals" with respect to 'conceptual organisation', 'spaciousness', 'extensiveness', 'new display' and 'instructional quality' derived

by factor analysis of a long list of defining features of the halls. Lakota claimed that these measures reflected the interests and wishes of the visitors and represented some basic design dimensions. As pointed out in the previous chapter, the exhibit design variables investigated by Shettel were the physical characteristics of the displays themselves, characteristics such as the readability level of labels, location of various displays and whether the displays contained moving parts. Neither Lakota nor Shettel were concerned with perceived exhibits.

Theoretical Outline

Conceptually, the present theory owes much to the ideas of the Gestalt psychologist Kurt Lewin (1890 - 1947) and, in particular, his notions of life space and valence.

A museum may be conceptualised psychologically as well as physically as the space in which a visitor moves. As a visitor moves physically through a museum he is also moving psychologically through his life space. Thus, his physical movements represent acts of psychological significance to him. According to Lewin, when a person is attracted by an object that object is said to have positive valence, an analogy he drew with the valence between atoms in chemistry. By contrast, negative valence refers to an object, or region in life space from which the person is repelled. Life space may contain regions with several valences active at once and these may give rise to conflict.

Different exhibitions will have different "valences" for different visitors. In the case of a visitor who has visited the museum before or other museums he believes to be similar, a good deal

of direct prior experience will influence the magnitude and directions of the valences of the different exhibitions. Consider, however, a visitor who has never visited the museum before or other museums he believes to be like it. Is it sensible in this instance to talk of exhibitions in the museum as having different valences when he has no experience of them? For any visitor, it is reasonable to suppose that topics or subjects dealt with in certain exhibitions will be of more interest to him than others; different topics can be said to have different valences. Thus, exhibitions will have attached to them valences that are quite independent of any direct experience of them. Of course, once a visitor has entered an exhibition, individual displays will be under the direct influence of the immediate perceptual process. However, all other things being equal, the valence attached to the visitor's interest in the subject matter will influence his perception of and behaviour towards the individual displays.

Individual displays in an exhibition can be thought of, psychologically, as occupying regions in a visitor's life space; and the present theory is primarily concerned with the perception of individual displays since it is these which attract visitors' attention. Each display will have a number of different valences active at once. Each of these valences will be attached to the characteristics or attributes of the display that are perceived by the visitor. For example, if the position of a display in the exhibition hall is a determinant of attractiveness as Melton's work has indicated, then 'position' will be a perceived characteristic of the display and it will have a positive or negative valence attached to it: the perceived position next to the exit of a gallery will have a negative valence. Similarly, the physical characteristics of displays, if they are perceived

by visitors will have valences that will influence the attractiveness of the display to a visitor. Thus, some perceived characteristics will have positive valences attached to them, some negative valences and other characteristics will conceivably be neutral, that is, characteristics will be graded on a continuum from positive to negative valence. The characteristics and the directions and magnitudes of the valences attached to them will vary from visitor to visitor.

A visitor will tend to be attracted to a display when the combinations of characteristics approaches the optimal combination for him. The optimal combination will result in the highest net positive valence when the valences are summed across all the valences attached to the perceived characteristics. As the combination of perceived characteristics moves away from the optimum and the magnitude of the net valence reduces, so the likelihood of him being attracted to the display decreases.

The valences attached to exhibits will also vary according to a visitor's proximity to them both physically and temporally. For example, a visitor setting off from home to the Natural History Museum to see a particular exhibit he has heard about might experience an increasing excitement in anticipation of seeing the exhibit, as he approaches the Museum. Once inside the Museum, he might become interested in other exhibits which confront him, and as a consequence his interest in seeing the exhibit which initially attracted him to the Museum might be diminished. Therefore, the valences of exhibits are not 'fixed' but will vary on a gradient from high to low according to circumstance.

The present theory is concerned with the attractiveness of exhibits as and when visitors experience them in competition with others in an exhibition. It is the strength of the exhibits' valences

at the time visitors experience them that will determine their attractiveness; and it is at this point that they will need to be measured.

The primary concern of the theory is with exhibits rather than individual visitors; and while both are different sides of the same coin, it is more helpful to focus on exhibits and the reasons why they attract or fail to attract members of a population of visitors. Thus, in the same way that a visitor was thought of being attracted to a display with the optimum combination of characteristics for him, an exhibit with the optimum combination of characteristics would be the one that attracted all visitors to an exhibition. Such an exhibit can be thought of as an ideal exhibit or an archetypical display; and real-world exhibits can be conceptualised in terms of their proximity to the ideal exhibit. A real-world exhibit close to the ideal will be one that is perceived to possess a large number of characteristics in common with the ideal, whereas an exhibit distant from the ideal will share very little in common. The proximity or distance of a real-world exhibit from the ideal is, of course, psychological distance, existing in a visitor's life-space and it has a much more precise meaning than the idea of distance in Lewin's theory.

The more characteristics a real-world shares in common with the ideal exhibit so the stronger will be its overall valence or attraction to visitors. Behaviourally, this means that a real-world exhibit which shares many of the characteristics associated with the ideal exhibit should have a high attracting power. Following on from this, it can be seen that the ideal exhibit will have an attracting power of unity and the measured attracting powers of real-world exhibits thus give a direct measure of their proximity to the ideal.

The theory thus predicts that the nearer the attracting power of a real-world exhibit is to unity, the more closely its perceived characteristics will resemble the perceived characteristics of the ideal exhibit. Therefore, once the characteristics of the ideal exhibit are known, the theory will provide an explanation of what makes an exhibit attractive, and conversely what makes one unattractive.

Testing the theory involves, operationally:-

1. measuring the attracting powers of real-world exhibits;
2. elucidating and measuring the perceived characteristics of the ideal and real-world exhibits;
3. computing the 'distances' between real-world exhibits and the ideal (in terms of their perceived characteristics);
4. calculating the correlation coefficient between the attracting powers of exhibits and their 'distances' from the ideal.

The remaining chapters of the thesis are concerned with testing the theory just outlined and subsequently with discussing the results and their implications for the design of museum exhibitions. Before turning to these issues, however, it is interesting to reflect on how the proposed approach resembles and differs from directly related fields of study investigated by psychologists.

The museum literature is replete with references to the 'curiosity value' of exhibits and it is usually claimed that successful exhibits are those that 'arouse' the curiosity in visitors. There is now, of course, a substantial literature concerning the study

of curiosity by psychologists (for a review see, for example, Maw and Maw, 1977), and its relevance to the present work seems obvious, at least intuitively. As pointed out by Maw and Maw, it is not clear from this literature whether curiosity refers to a state brought on by certain kinds of stimuli or whether it refers to a trait or individual disposition.

Typically, experimental investigations of curiosity have involved exposing subjects to objects of one sort or another and subsequently asking them about their preferences as measures of curiosity or alternatively measuring aspects of their behaviour, for example, the amount of time spent looking at different objects, facial expressions, etc. Explanations of why certain objects arouse curiosity rather than others is usually given by reference to the complexity or familiarity of the stimulus objects and there is a plethora of reports devoted to studies in which the nature of the stimulus objects is manipulated and treated as a source of independent variables. A similar approach is often adopted in the study of aesthetics or, more generally, studies of familiarity and liking (Sluckin, et al, 1982).

The 'trait' or personality approach to studying curiosity has usually involved administering a battery of tests to measure abstruse notions such as Epistemic curiosity or sensation-seeking curiosity (Maw and Maw, 1977) which are then related to preferences for certain objects or ideas. A similar approach has arisen within environmental psychology with the idea of environmental dispositions, that is, individual differences in personality as it functions in relation to the everyday physical environment (Craik 1970, 1976). Researchers in both these traditions seem to be asking essentially the same question, namely, is it possible to distinguish between individuals' attraction to objects in terms of

certain immutable personality traits?

The 'arousal' view of curiosity bears some resemblance to the theory of exhibit effectiveness outlined earlier in the chapter. Both approaches are ostensibly concerned with discovering why certain objects are attractive and why others are not. However, the affinity ends here. The proposed theory seeks an explanation in terms of the cognitive constructions placed upon objects by individuals whereas investigators into the 'curiosity-value' of objects have sought explanations in terms of the moribund 'stimulus-response' paradigm of American empirical psychology. It is proposed that 'curiosity' should be redefined in the terms of the present theory.

The notions of 'curiosity traits' and 'environmental dispositions' are so impoverished, both conceptually and theoretically, that there are no grounds to suppose that any of the so-called traits or dispositions measured and reported in the literature have any bearing on the present issues. However, these considerations should not be taken to imply that there are no individual differences between visitors in terms of their attraction to particular displays; clearly there are differences as not all visitors are attracted to all displays! For the moment, perhaps all that needs to be said about this from the point of view of the present theory, are that the primary determinants of the attractiveness of the display are the valences of the perceived characteristics of the display. Other variables will only affect attractiveness if they firstly influence an individual's perception of the display. (The issue and importance of individual differences to the theory will be taken up more fully in Chapters 7 and 8 of the thesis.)

In summary, the chapter has outlined a new theory of exhibit effectiveness in which real-world exhibits have been conceptualised

in terms of their proximity to an ideal exhibit. The ideal exhibit is the one that attracts all members in a population, that is, has an attracting power of unity. Real-world exhibits close to the ideal will have high attracting powers and will share many of the characteristics visitors perceive in the ideal exhibit. As an exhibit's attracting power decreases, so will the number of perceived characteristics it shares in common with the ideal decrease.

To test the theory it is necessary to:-

- (i) measure the attracting powers of exhibits;
- (ii) elucidate and measure their characteristics (and the characteristics of the ideal exhibit) as perceived by visitors;
- (iii) compute the distances between real-world exhibits and the ideal, and
- (iv) calculate the correlation coefficient between attracting power and distance from the ideal.

Chapters 4 and 5 are concerned with developing a reliable method for measuring the attracting powers of exhibits in the Hall of Human Biology and Chapters 6 and 7 are concerned primarily with collecting data with which to test the theory.

CHAPTER 4

METHODOLOGICAL APPROACH TO MEASURING VISITORS' BEHAVIOUR IN THE HALL OF HUMAN BIOLOGY

General approach to measuring the attracting and holding power of museum exhibits.

Webb et al (1966) cite two examples of the application of erosion and accretion measures to measuring the attractiveness of museum exhibits:

"The relative popularity of exhibits with glass fronts could be compared by examining the number of noseprints deposited on the glass each day (or some sample of time, day, month, and so forth). This requires that the glass be dusted for noseprints each night and then wiped clean for the next day's viewers to smudge. The noseprint measure has fewer content restrictions than most of the trace techniques, for the age of the viewers can be estimated as well as the total number of prints on each exhibit. Age is determined by plotting a frequency distribution of the heights of the smudges from the floor, and relating these data to normative heights by age (minus, of course, the nose-to-top-of-head correction)."

"A committee was formed to set up a psychological exhibit at Chicago's Museum of Science and Industry. The committee learned that the vinyl tiles around the exhibit containing live, hatched chicks had to be replaced every six weeks or so; tiles in other areas of the museum went on for years without replacement. A comparative study of the rate of tile replacement around the various museum exhibits could

give a rough ordering of the popularity of the exhibits. Note that although erosion is the measure, the knowledge of the erosion rate comes from a check of the records of the museum's maintenance department."

A more precise method of measuring the physical traces made by visitors in an exhibition was carried out by Bechtel (1967) using an ingenious device known as a hodometer which was installed under the floor covering. A hodometer is an interconnecting network of electrically wired pads that records the location and number of footsteps in a gallery.

Physical trace measures of the sorts just described do not permit one to make generalised statements about the relative attractiveness of all the exhibits among a sample of visitors nor do they allow a comparison of the attraction of different exhibits to an individual visitor. These data can only be obtained by tracking visitors as they go round an exhibition, noting which exhibits a visitor looks at, and the time he spends at each one. In addition, by following visitors, it is possible to note the order in which a visitor looks at the exhibits and subsequently plot the routes taken through the exhibition as a whole. In fact, typical observational studies in the museum setting have been carried out by an observer physically following visitors through an exhibition. A fixed camera has been used instead of an observer when the exhibition was concentrated in a fairly small space (Nielson 1946). However, at crowded times it may prove difficult to identify and trace the course taken by selected individuals (Nedzel 1952).

Early pilot studies in the Hall of Human Biology were undertaken in which the observer physically followed visitors along the lines of the studies reported in Chapter 2. However, in nearly all

attempts at tracking individuals unobtrusively by this method, before very long in the course of an observation, the visitor became aware that somebody was watching him. Apart from the embarrassment to the observer upon being discovered, the realisation by the visitor that he was being followed had a disturbing effect on his behaviour. It was quite noticeable that the behaviour of visitors after they had discovered they were being watched was different from their behaviour before. Although there is contradictory evidence in the literature on the extent of this sort of bias in naturalistic research in general (see Hollenbeck, 1977) there was no doubt about the biasing effects of the intrusiveness of the observer upon the behaviour of visitors in the pilot studies. It is interesting to note that other researchers do not appear to have had similar difficulties. For example, Sheppard (1960) reported that only 3 per cent of visitors observed at agricultural exhibitions had claimed they were aware of being observed when questioned at the end of their visits. Reported studies of Museum visitors do not generally mention the problem of observer intrusiveness although Melton (1935) was an exception. He pointed out that it was not possible for visitors to be tracked during quiet periods as they tended to become aware they were being followed.

Specific methodology adopted in the present study

In the light of the foregoing considerations it was decided to develop a closed-circuit television (CCTV) system which would permit the movements of individual visitors to be tracked unobtrusively as they wandered through the Hall of Human Biology. This system was developed after investigating closed circuit installations in twenty other institutions in the London area. The list of institutions included department stores as well as other museums.

Closed-circuit television system and its operation

The present set-up in the Hall of Human Biology is as follows:

There are twenty six cameras located throughout the sections in the Hall of Human Biology. Twenty cameras operate in areas with natural or good ambient lighting and are fitted with videon tubes. Six cameras are situated in low lighting areas and are fitted with an added lens unit. All cameras carry lenses with angles of view and focal lengths which optimise their fields of vision without compromising picture acuity on the television monitors. With the exception of one camera in Section I of the exhibition (see Chapter 1) which is fitted with a pan unit, all the other cameras are in fixed position.

The cameras are relayed to one of two monitors situated in a control room adjacent to the exhibition. One of the monitors is connected to a video-recorder (National Time Lapse VTR 8030). The layout of the monitors and controls are shown in Figure 3. Below the monitors is a special keyboard used for logging visitors' behaviour (marked B in Figure 3) which was developed after the CCTV system was installed and whose function will be described in detail in the next section. To the right of B are the master controls for switching the system on and off, for selecting which of the TV monitors to view and the on/off controls for the video-recorder (VCR). The area marked C contains controls for selecting the different operating options of the VCR equipment, iris controls for the cameras operating in low light areas and the pan unit switch for the camera in Section I.

The observer is seated facing the television monitors. Once the system has been turned on and the monitor/s to be viewed and the

VCR options selected, the observer is able to track the movements of an individual visitor by switching between cameras as appropriate and by viewing the camera images relayed to the TV monitor. Camera selection is effected by the observer through the operation of camera switches which are situated on the main control panel, marked A in Figure 3. Figure 4 shows in more detail the layout of the control panel A; the control panel is isomorphic with the ground plan of the exhibition. The switches are denoted by 'CCTV camera positions' in the figure. (As well as indicating the direction of the cameras they are switches for selecting which camera view is relayed to the monitor). Apart from minor adjustments to the iris controls and operation of the pan unit as necessary, the observer is concerned solely with interacting with the Figure 4 control panel during an observation.

Once an observer is familiar with the layout of the exhibition, it requires less than a day's practice to become proficient at following visitors through the exhibition using the CCTV system.

There are, of course, ethical questions in observing the behaviour of museum visitors unobtrusively, the more so when visual records are kept without the visitors' permission. In recognition of this, notices were placed throughout the exhibition with the express purpose of informing visitors that closed-circuit television was in operation to monitor the effectiveness of exhibits and also for security purposes.

A pilot study to evaluate the unobtrusiveness of the system was undertaken in which 18 visitors were interviewed upon entering the exhibition, then they were followed during the course of their visits using the CCTV system and finally interviewed again upon leaving. None of the visitors interviewed was aware of being

followed and upon being informed of the fact, none objected. (Somewhat disturbingly, all of them accepted that surveillance by CCTV was part and parcel of modern life and they appeared to be inured to its intrusiveness. These findings do not, of course, remove the ethical questions but they do indicate that the researcher might be more sensitive to the ethical issues than the general public.)

Behavioural recording system

The observation of a visitor's behaviour and the subsequent analysis of it is a time-consuming business and can take as long as three hours or more for a single subject. This fact together with the dubious ethic of keeping permanent video recordings of visitors' behaviour, led to the development of a real-time recording system which allows the observer to measure the times a visitor spends viewing exhibits while also monitoring his behaviour using the CCTV.

Before considering the important structural and functional aspects of the behavioural recording system, the nature of the observer's task in operating it will be described.

To recap, the observer switches between cameras by interacting with control panel A shown in Figure 4. The alphanumeric codes on control panel A in Figure 4 refer to the individual sections or assemblies in the exhibition. The numerical codes, 1, n, in each section refer to the exhibits or displays in each assembly. This information serves as an aide-memoir to the observer as he or she logs viewing times on the special keyboard (marked B in Figure 3).

The recording system acts basically as a sophisticated stopwatch

Image redacted due to third party rights or other legal issues



Image redacted due to third party rights or other legal issues



with the ability to record, process and replay its findings. During logging, the system provides a real-time display of times and status. These data are recorded along with information to identify the visitor and his activities. During replay of this information, the results are presented on the computer's visual display unit at a rate controlled by the operator. An important consideration during the design of the system was that it should require a minimum of operator training and a minimum of technical help during day to day use. To these ends, the main programme is arranged to operate conversationally. That is to say, it prompts the operator for information it needs, and it attempts to detect operator errors without wasting data and without grinding to a complete halt. The main elements of the hardware will now be described.

Hardware

The hardware is divided functionally into six components as shown in Figure 5. The solid arrows indicate directions of data flow during normal operation. The dashed arrows indicate possible directions of flow but these are not normally used. The components within the outline box are housed in a Commodore PET micro-computer which forms the basis of the hardware of the system.

The central processor unit (CPU) is the heart of the computer proper and it is based on a type 6502 microprocessor. This processor uses 8-bit data bytes and can address up to 64K bytes directly (K = 1024 in this context). 32K bytes of random access memory are available for programme storage and execution of the user's programme. The rest of the memory space contains the interpreter to execute the user programme; subroutines for input, output and peripheral control and the monitor programme which supervises the overall operation of the computer. The visual display

unit (VDU) is a TV-type screen which displays continuously the data stored in code form in a particular area of memory as graphical symbols. The VDU has three main functions in the recording system, viz.;

- (i) to pass information and questions to the operator during the setting-up of observation or data reading sessions;
- (ii) to display real-time results and status during observational data recording sessions, and
- (iii) to display previously recorded results during data read-outs.

The main keyboard is normally used only in setting-up and during the data read-out phases of operation. The keyboard is a full alphanumeric type arranged approximately as a standard typewriter plus an additional numeric keypad. There are also keys to allow the operator to delete, change or insert characters so that typing errors in replies can be corrected before the computer acts on them.

The keyboard is an integral part of the PET computer and cannot function independently of the CPU. This is advantageous in the present application in that, being interfaced directly to the CPU, the keyboard is responded to much faster than in many systems. (Time-sharing systems are particularly poor in this respect; and the typical time-sharing delays would render an application such as this quite impracticable.)

The main keyboard is quite unsuitable for use by the observer to log visitors' behaviour; it is too cluttered, it is densely laid out and it requires the observer to enter information into the

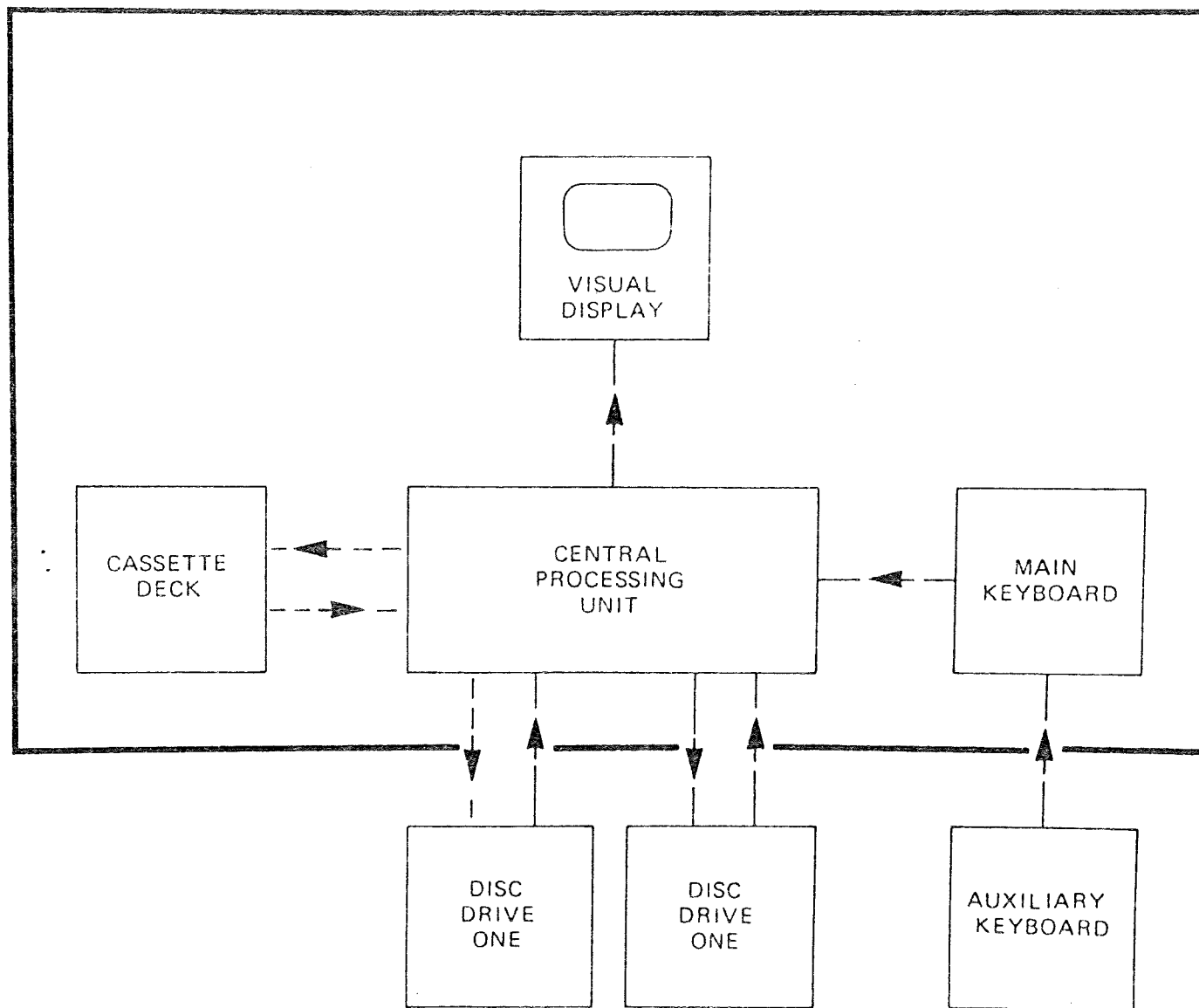


Fig. 5: Functional layout of real-time recording system. (The components within the outline box are housed in a Commodore PET microcomputer).

the computer at real-times rates. Therefore, a special keyboard (B in Figure 3) was built for recording visitors' behaviour. The keys used are high quality, touch sensitive type set on an 18 mm matrix as used on large electric typewriters. They are physically arranged in accordance with their logical functions. The layout of the keyboard and the functions of the keys will be described in more detail later.

In many systems, an auxiliary keyboard would be interfaced to the CPU as an extra input port to the computer. In the present system, the keyboard is simply connected in series with the main keyboard of the PET computer shown in Figure 5. The CPU is thus unaware that an external keyboard is being used. This arrangement retains all the speed advantages of the integral keyboard and also simplifies the hardware and software requirements. The system can accept data from the keyboard at up to fifty characters per second and can process up to six discrete events per second. These rates are considerably faster than required for the present application.

The random access memory in the CPU is capable of high-speed use (access typically within half a microsecond), and it is useful for short term storage of moderate quantities of data. However, this memory is volatile, that is, its contents are lost forever whenever the system power supply is turned off. It is also too small to store the quantities of data generated during a typical observation. The most suitable medium to do this job is the magnetic disc. The discs used in this system are the smallest common variety and are known as 'mini-floppy discs'. The disc is housed in a unit known as a disc drive. The drive rotates the disc at a steady speed and positions a magnetic reading/recording head over the disc in response to CPU commands. The system used

here houses two drives in one case, physically separate from the main computer housing. Special purpose interfacing software and hardware were installed within the PET to control the disc units. The drives allow the storage of approximately 105K bytes of data per second.

Drive one is used to store the main survey programme only. The first thing the observer does after switching on the system is to load the main programme from this disc drive. The drive is set to read only during normal operation, removing the risk of accidental erasure or alteration by operator errors. Only about one-eighth of this disc drive's capacity is used to store the programme and under normal conditions it is used only once at the start of each working day.

Drive two is used to store the data generated during the observations, and to replay these data as and when required. It is therefore set to be able to read or write discs and the programme has to be careful to prevent the operator accidentally erasing or recording over previously recorded results.

The PET also has an integral cassette deck for non-volatile, bulk data storage on standard compact cassettes. However, apart from some early trials, the cassette deck was not used with the behavioural recording system as it proved far too slow and unreliable. However, it was used to keep back-up copies of the discs at various stages as an additional safeguard against the loss or accidental damage of a disc which is physically less robust than a compact cassette.

One other aspect of the PET's hardware is of particular relevance to this application. The PET contains two real-time clocks, one

dependent on the other. The independent clock is a simple, two-byte, integer count that is incremented once every sixtieth of a second. The dependent clock forms an hours, minutes, seconds representation of the integer count. (The latter clock is used only to provide a clock-time display for operator convenience in the present application.) The clocks are quartz crystal controlled and although the crystal is untrimmed it is sufficiently accurate for this work, the particular sample being used had an accuracy of about 0.026%, that is, an absolute error of approximately 0.02 seconds in every 100 seconds. This magnitude of error is insignificant compared with other external experimental errors such as operator reaction time.

Outline of software and operation

The programme is written in the BASIC language and occupies approximately 10K bytes of memory. It requires a further 4K bytes during execution. A permanently stored interpreter converts the programme instructions to machine code for execution by the CPU. Input and output interfacing is handled by a permanently stored monitor programme. After switching on the machines, the operator must instruct the computer to load and run the behavioural recording programme which is stored on Drive one. This requires three simple commands and takes a few seconds. From this point on, the programme questions and prompts the operator when it needs information or a decision.

When the programme is started for the first time, it asks for the time of day and the date. It then sets an internal clock for reference until it is switched off again.

The operator is then offered three choices:

1. To start a new session of observational recording.
2. To read out data previously recorded.
3. To terminate the session.

The selection is phrased as a multiple-choice with a reply of 1, 2, or 3. If the operator chooses '3', the system simply displays a reminder to switch off the power supply before leaving and hands control of the computer back to the monitor. The other two choices will now be considered in some detail.

If the operator selects choice '1', he wishes to log the behaviour of a visitor in real-time. Before describing the operating procedures, the special auxiliary keyboard used for logging behaviour will be discussed in more detail, as it is central to the control of this part of the programme.

Figure 6 shows the layout and labelling of the auxiliary keyboard.

NEW VISITOR

This is used to tell the programme that the operator wants to start tracking a new visitor. Any tracking in progress is terminated and the timers are reset.

V. AT EXHIBIT

This is used to indicate that the visitor has now stopped at a display. A timer is reset and starts to time the visitor at this display. The programme now expects a code number for that display to be entered before any new activity is signalled by the operator.

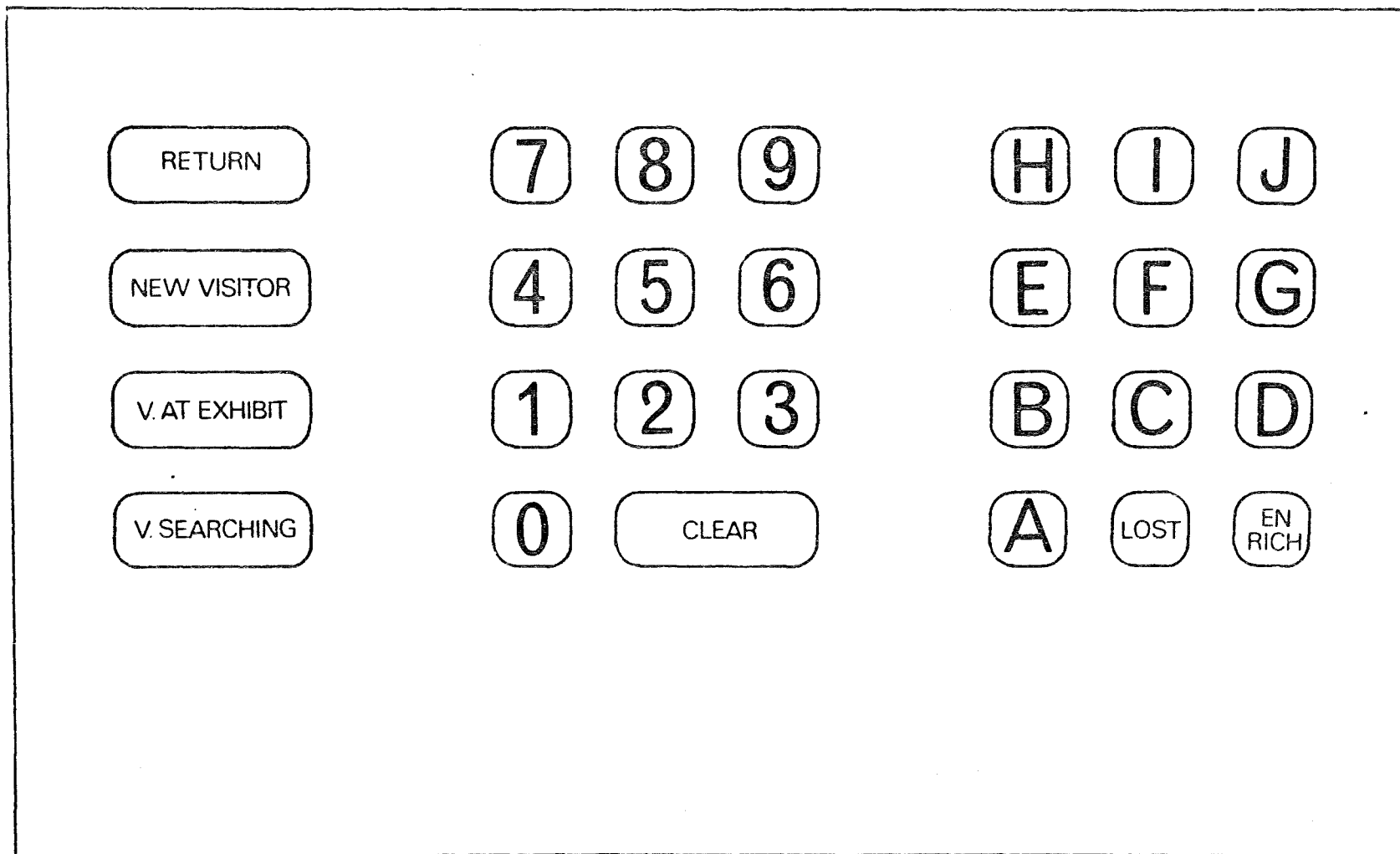


Fig. 6: Layout of auxiliary keyboard used for logging visitors' behaviour.

V. SEARCHING

This is used to indicate that the visitor is moving around rather than stopped at any particular display. A timer is reset and starts to time this activity. No further information is expected until the start of some new action.

Letters A to J. These are used to indicate which assembly of the Hall of Human Biology the visitor is in. The relevant key is pressed at the moment the visitor enters the assembly. A timer is reset and starts to measure the total time elapsed in that assembly. The visitor is assumed to be searching initially, so another timer is reset and starts to time the elapsed search time.

Numbers 0 to 9. These are used to enter the code number of particular displays just after pressing V. AT EXHIBIT. They may be used also during other phases of programme operation to answer multiple-choice questions and enter observation serial numbers.

CLEAR

This may be used to correct erroneous assembly letter or display number entries. During a searching period this key will clear the assembly letter for re-entry. During an 'at display' period, it will clear the display number.

ENRICH

This is used just after an exhibit letter key to indicate that the visitor has entered the subject enrichment area (see Chapter 1) associated with that assembly. The letter 'E' is appended to the indicated assembly letter in the display.

LOST

This is used to indicate to the computer that the operator has lost visual contact with the visitor. The time elapsed in this state is measured, and the last known location of the visitor is noted. If contact is re-established, the operator presses the assembly letter in which the visitor reappears.

RETURN

This key is used when a reply to the question posed by the computer is more than one character long. It is used simply to indicate the end of the reply. When the computer is expecting a one character input only, this key is not needed and the first character to be pressed is taken as the reply.

Operating procedures for recording a visitor's behaviour

The computer begins by asking the operator to 'hit any key to open on disc drive two'. This is simply a safeguard to allow the operator to abort the programme without risk to the disc (by pressing STOP) if he realises that the wrong disc is in the drive or the wrong response was given in the initial choice of options. If the operator presses anything other than STOP, a new file will be set up on the disc and an appropriate entry made in its index. The file will receive the data from the observations. All such files require names for unique identification in the future. The programme forms a suitable name by combining the letters SURV with the time and date at the moment of opening the file. For example, the file opened at 09.58 hrs on 4th July, 1979 would have the name 'SURV095804JUL79'. This method of naming ensures unique

names and allows the operator to tell when a file was created just from its name.

When the file is opened, the operator is told its name and asked to note it, if required, for future reference. The operator is then asked to 'press any key to continue'. This is a device used at many points throughout the programme to allow the operator to pause or perform any other task such as noting down information from the display.

When the operator presses to continue, the computer asks for an observation number to identify the visitor about to be tracked. When this is given, the operator waits for the visitor to pass a predetermined entry point and presses the letter key for the exhibit the visitor has entered. All the internal timers are reset, and the clock time at which the visitor entered is recorded. The VDU now sets-up a standard layout of information used through active information logging. The programme assumes that the visitor is initially searching at the point of entry.

The operator then uses the special keyboard to log the visitor's activity and location. For example, if the visitor stops to look at display 2, the operator presses the 'V. AT EXHIBIT' key followed by the '2' key. The lower section of the display now changes to reflect the new situation. The time elapsed in the assembly continues to count, but the lower line changes to show the display number and the time elapsed at this display.

Similarly, when the visitor enters a new assembly, the operator presses the letter key corresponding to that assembly. The assembly letter in the lower lines of the display now changes and the visitor is assumed to be searching again. Both elapsed time counters are

reset to zero, although the overall time elapsed in the upper part of the display continues to count.

The operator continues logging the visitor's activities until the visitor leaves the exhibition. The VDU continues to show the current state of affairs although in practice the operator has no opportunity of viewing the VDU while he is logging viewing times. When the visitor leaves the exhibition, the operator presses the 'NEW VISITOR' key. The final items on the last visitor are recorded and the observation is terminated.

Other operating procedures

The computer now asks the operator if he wants to go on to make another observation or if he wishes to stop. If the operator wishes to start another observation, the computer in effect asks the operator to open a new file and the process just described is repeated. If the operator chooses not to start another observation, the operator is again offered the three choices that feature in the programme. If the operator selects choice '2', he now wishes to read out previously recorded information from disc. A large part of the programme is concerned with assisting the operator to find and identify the file required without causing errors which might stop the programme. For example, in the simplest way of operating the disc drives, if the operator requested a file by a name which did not exist on disc, the system would stop the programme, report the error, and hand control over to the computer monitor. This would waste a lot of time. However, avoiding this clumsy procedure does incur penalties. When the choice is made to read out data, there is an enforced delay of about 15 secs. while the programme assembles its own directory of all the files present on disc. This directory is needed to allow the programme to

establish whether or not a requested file exists. The directory is also used to tell the operator the names of all the files on the disc to assist in the identification of the requested file. The operator is asked to wait while the directory is prepared. The programme then asks if the name of the required file is known. If the operator does not know the name of the file, the computer asks if an explanation of the name finding routine is required. If it is then it is given. All the names of the files on the disc are presented to the operator, one at a time. The computer asks if each one is the one required. If the operator presses 'N', the name of the next file is presented, and so on until the operator presses 'Y'. When the operator has chosen a file name, the programme asks for confirmation that reading from this file is to go ahead. (This is to allow the operator to use the name finding routine alone just to check file names and to allow the correction of keying errors.) The programme then checks the directory, if necessary, to confirm the existence of the file. If it is not present, the operator is informed and execution returns to the stage where the operator is asked if the name of the required file is known. Read-out from the file then starts. The information is read in blocks which almost fill the screen. The operator is then invited to press a key to continue to the next block. Successive blocks 'scroll-up' on to the screen in conventional fashion.

The information is 'unpacked' from the condensed form recorded on the disc, and turned into a more literal style to improve its readability. Clear delineation is made at the start of new observations and between the passage of the visitor from one assembly to another. When the visitor passes into a new assembly, total elapsed times for that assembly are shown - one for the total time spent searching, and one for the total at displays. To avoid any rounding errors, these times are recorded explicitly by the data

logging programme rather than being produced by individual summation in the read-out programme. Similar totals are given for the entire observation when the visitor leaves the exhibition. The read-out programme would be stopped if any attempt was made to read data beyond the end of the recorded file.

To summarise, this chapter has described a novel system which was developed specially to observe and record the behaviour of visitors to the Hall of Human Biology. Although tailor-made for collecting data pertinent to the computation of measures of the attracting powers and holding powers of exhibits, the system could be easily adapted to measure other aspects of visitors' behaviour. The real-time recording system could function, generally, as an event-recorder; as the function of the keys of the auxiliary keyboard (Figure 6) are purely nominal, they could easily be taken to signify different events.

In Chapter 5 the reliability of observations of museum visitors will be discussed in general, followed by a study of inter-observer reliability using the system described in this chapter.

CHAPTER 5.

PILOT STUDIES OF SYSTEM RELIABILITY

The Reliability of Observational Methods in Museum Research

Reliability as a concept incorporates the separate components of accuracy and stability (Kerlinger, 1964). Accuracy is concerned with precision, that is, whether the instrument is a precise representation of what is being observed. Stability is concerned with the consistency of the results obtained when repeated measures are made with the same or similar instruments under similar conditions. In the present context, a reliable method of observing visitors is one that yields stable, consistent and dependable estimates of attracting power and holding power of all exhibits, with only small errors in measurement, when different samples of visitors are observed by different observers.

General problems of reliability in observation research

Apart from the problem of observer intrusiveness mentioned in the previous chapter, there are three possible types of error when different observers set out to collect the raw data necessary for the computations of attracting power and holding power on any one occasion. Firstly, there is a type of error which may be characterised as an 'error of commission', that is, recording a visitor as looking at an exhibit when, in fact, he or she is not. Secondly, there is a complementary 'error of omission' which would be failing to record a visitor as looking at an exhibit when in fact he or she is looking at it. Thirdly, there is the possibility of observers having agreed that a visitor has stopped at an exhibit, subsequently disagreeing about the amount of time the visitor spends viewing the exhibit. As well as the possibility

of these types of errors occurring among different observers when observing a single visitor at a single exhibit, it is also possible that the extent of these occurrences will vary from visitor to visitor and exhibit to exhibit. Also, there is the problem of observer drift; this involves stability over repeated observational measurement times and it has been labelled rather unprepossessingly in psychometric theory as "observer decay". In museum studies, the term 'decay' implies that an observer is less accurate at subsequent times than he or she is initially.

Mitchell (1979) has pointed out that there are three different coefficients which when computed purport to reflect the reliability of observational data; namely, an interobserver agreement percentage, the reliability coefficient and the generalisability coefficient.

In the present context, the number of agreements and disagreements between two observers about whether a sample of visitors stopped or did not stop at the various exhibits they encountered would yield an easily computed interobserver agreement percentage on stopping behaviour. However, such a percentage would not indicate how interobserver agreement varied between different visitors and different exhibits over time.

An interobserver reliability coefficient would be obtained, for example, from the correlation between two observers' estimates of the times individuals in a sample of visitors spent looking at various exhibits. The 'true score' implied by the correlation would reflect real differences in holding times among the sample of visitors and the 'error' would reflect differences between the observers in their estimates of the time spent at exhibits along with random error. A reliability coefficient computed in this

way will be lower when a homogeneous group of visitors has been observed than when a more heterogeneous group is observed because in the former case there is a low true score variance and in the latter case a high true score variance.

In the examples just given it is possible to have high interobserver percentage agreement but at the same time to have a low reliability coefficient. For example, two observers might have a very high agreement about the exhibits a sample of visitors stopped at while at the same time disagreeing as to the actual duration of time individuals spent viewing exhibits. This serves to illustrate that neither observer agreement nor reliability coefficients by themselves adequately describe the reliability of measuring instruments used in observational studies.

One way of approaching reliability is, therefore, to compute a variety of interobserver agreement percentages and reliability coefficients appropriate to the measures being taken and to look for consistency across all the coefficients. However, an alternative approach has been recommended by Mitchell (1979); and this is to think of observer reliability in terms of an analysis of variance model or Generalisability theory as it has now become known, (Cronbach et al, 1972). In this approach, different sources of variation - for example, different observers, different visitors, different exhibits and different observation periods - are known as different facets. In a generalisability study (G study) the idea is to estimate the contributions or variance components of each of the facets to the overall variation in the set of observational measures.

Mitchell (1979) has argued that this approach permits a researcher to look at the proportion of variance in scores that is attributable to the consistent behaviour of the observer and she has shown that this

can be higher than the component of variance attributable to subjects (Mitchell 1979). However, the idea that overall error can be attributed to different facets, and their different components of variance estimated is rather optimistic. The only things which may be directly estimated are the variances of combinations of errors. All indirect estimates rely on some sort of additive model for the errors. For example, in an inter-observer reliability study with two observers, the 'between-observer' variation is estimated by subtracting an estimate of the 'within-observer' variation from an estimate of the total variance. The problem in this example is that the estimate of 'within-observer' variation is extremely poor as it is based on only one degree of freedom. On the other hand, the estimate of the total variance is likely to be based on a much larger number of degrees of freedom and is likely to be quite good. In subtracting quantities of radically different orders of precision it is even quite easy to obtain a negative value for the estimate of 'between-observer' variance. The only way to improve matters is to increase the number of observers but this is not always possible in projects where manpower is strictly limited. Another problem arises when the numbers of units in each of the facets are unequal because analysis of variance methods are not very well developed to deal with such instances; and the problems increase as the number of facets increases. In a carefully controlled, properly balanced laboratory study with a normally distributed response, components of variance might be useful but in most naturalistic and applied projects it is much more sensible to report the variation of measurable quantities, avoiding components of variance altogether.

Reliability studies of CCTV and behavioural recording system.

In assessing the reliability of the behavioural recording system,

the present study was designed to investigate two potential sources of unreliability. These were

- (i) errors in recording visitors' behaviour and,
- (ii) observer drift.

Two female observers were available for the main study (reported in the next chapter) and they also participated in the reliability studies, one of which was conducted at the commencement of the main project and the other at the end. These studies were concerned mainly with obtaining measures of the extent of interobserver agreement; and in line with the comments made in the previous section, G theory was not used.

The two assistants were trained in the use of the CCTV system and the behavioural recording system prior to the commencement of the reliability studies; and the data collected during the studies formed part of the total data collected in the major study reported in the following chapter.

Subjects

In each study fifteen subjects were systematically selected as they entered the Hall of Human Biology (see Appendix 1) and subsequently they unknowingly participated in the studies.

Method

In both reliability studies, the fifteen visitors selected from visitors to the Hall of Human Biology (this number represented ten per cent of visitors participating in the main study) were observed using the closed-circuit television and the behavioural recording system. Video-recordings of these observations were

kept. Each assistant took it in turns to make observations; thus one assistant made seven and the other eight recordings in each study. Each assistant was instructed to use the same criterion in judging whether or not a visitor was 'at an exhibit'. Visitors were to be judged 'at an exhibit' when their gaze appeared to be directed at the information presented in an exhibit, regardless of whether they were stationary or moving. Fleeting, side-ways glances at exhibits were not recorded. (Such glances are too short to record and hardly constitute 'looking at' an exhibit.) This definition of 'at an exhibit' was different from one used successfully by the author (Alt, 1979) in a study of a temporary exhibit elsewhere in the Museum, when visitors were only recorded as looking at an exhibit if they stopped in their tracks. This criterion is unrealistic in the Hall of Human Biology where certain exhibits - graphic panels strung along a wall, for example - necessitate movement by the visitor if he is to look at the entire exhibit. It is of crucial importance for observers to apply the same meaningful standards if the reliability between observers is to be measured in terms of their agreement. Herbert and Attridge (1975) have suggested that agreement between observers is not necessarily a measurement of reliability; agreeing observers could be applying the same incorrect behavioural definitions. However, it seems they have missed the point they were trying to make. Two observers applying the same (incorrect) definitions could be very reliable but their measurements would be invalid.

After making an observation, each observer analysed the data she had recorded using the read-out procedures described in the previous chapter. She made a listing of all exhibits at which she had recorded the visitor viewing and not viewing as well as the lengths of time she had recorded them viewing. Each assistant then analysed the video-recordings of the visitors they had not

observed themselves using the behavioural recording system and preparing a listing similar in format to the listing compiled by the other assistant. In each reliability study, comparisons were made between the listings generated for each observation by the two different methods.

Results

During the course of the first reliability study a fault developed in the hardware in the control panel of the closed circuit television system which affected the quality of recording on the video-recorder. As a result, four of the observations in the study had to be discarded. This fault was not rectified satisfactorily before the commencement of the second study and five recordings in the second reliability study were discarded for similar reasons. In all instances, there was a pronounced picture flicker on the recordings which made it impossible to follow the movements of the visitor being observed.

Disagreements over recordings in stopping behaviour

It is possible that some visitors might be more difficult to observe than others either because of the speed with which they move or the level of crowding in the exhibition. For this reason, each observation has been treated separately rather than aggregating the data across all visitors and making a single comparison in both studies.

For two observers, there are four possible outcomes in recording a visitor's behaviour at an exhibit; these may be cast in a familiar two-way contingency table as follows: (see Footnote on page 78).

		<u>Observer A</u>		
		Not stopped	Stopped	
<u>Observer B</u>	Stopped	<u>a</u>	<u>b</u>	<u>a + b</u>
	Not stopped	<u>c</u>	<u>d</u>	<u>c + d</u>
		<u>a + c</u>	<u>b + d</u>	<u>N</u>

The values of a, b, c and d for each observation in both reliability studies are given in Table 1.

Footnote:

Exhibits included in the analysis constituted only those exhibits along the routes taken by visitors during the course of their visits. Clearly, in the whole exhibition, for any visitor there may be some exhibits that he does not encounter simply because they are not on his route.

Table 1

Agreements/Disagreements about Stopping Behaviour between Observers
in Visitors' Stopping Behaviour. (Definitions of symbols are
provided in the text).

Observation Number.	1st Reliability Study				2nd Reliability Study			
	<u>a</u>	<u>b</u>	<u>c</u>	<u>d</u>	<u>a</u>	<u>b</u>	<u>c</u>	<u>d</u>
<u>1</u>	0	42	18	3	3	29	62	2
<u>2</u>	0	22	64	3	2	26	42	7
<u>3</u>	3	22	38	4	4	29	42	2
<u>4</u>	3	30	12	3	2	38	13	3
<u>5</u>	3	19	38	3	2	10	22	1
<u>6</u>	1	10	11	1	1	44	71	2
<u>7</u>	1	23	21	2	1	18	28	1
<u>8</u>	2	28	54	4	1	16	32	5
<u>9</u>	2	39	48	3	9	38	31	1
<u>10</u>	0	10	33	3	3	41	38	2
<u>11</u>	1	13	40	0	-	-	-	-

The percentage agreements between observers may be calculated easily from the ratio $(b + c)/N$ and these are presented in Table 2. On this basis it can be seen that the interobserver reliability is reasonable since they are in agreement in all instances for approximately ninety per cent of the time. However, the 'percentage agreement' measure is rather a blunt instrument for it gives no information about the nature of disagreements between observers. The same argument applies equally to measures of correlation such as the tetrachoric correlation coefficient (McNemar, 1962) which if applied to these data, would certainly show a high correlation between observers. Even though there is good agreement between observers, it may be that bias exists between them. For example, does one of the observers consistently record visitors stopping at exhibits while the other does not? An appropriate test for this sort of bias is given by the Sign Test (see Siegel, 1956) where the null hypothesis is that for any event about which the two observers disagree, the probabilities of the two kinds of disagreements are equal. This test can also be achieved by applying the formula (see Maxwell, 1961):

$$Z = \frac{|a - d| - 1}{\sqrt{a + d}} \quad ; \quad \text{Applying this formula and referring the}$$

results for all observations to the normal distribution, it may be concluded that any apparent differences between observers are likely to have occurred by chance (all $p_s > .10$).

Difference between observers in their recording of exhibit viewing times.

Any exhibit, \underline{i} at which observer \underline{A} and observer \underline{B} agreed a visitor had stopped will have two associated viewing times \underline{A}_i and \underline{B}_i , and the difference between them, \underline{d}_i , is given by $\underline{A}_i - \underline{B}_i$. In all

Table 2

Inter-observer Agreement Percentages in Visitors' Stopping Behaviour

	1st Reliability Study	2nd Reliability Study
Observation Number	percentage agreement between observers	percentage agreement between observers
<u>1</u>	95	95
<u>2</u>	97	88
<u>3</u>	90	92
<u>4</u>	88	89
<u>5</u>	90	91
<u>6</u>	91	97
<u>7</u>	94	96
<u>8</u>	93	89
<u>9</u>	95	86
<u>10</u>	93	94
<u>11</u>	98	-
Average	93	92
<u>S.E.</u>	1	1

observations, the values of the majority of d_i s were typically small with a few high outlying values. The variations in the distributions of d_i s are summarised in Table 3 which shows the inter-quartile ranges for each observation in both reliability studies.

Table 3

Interquartile ranges of distributions of d_i s for each observation in both reliability studies.

Observation Number.	1st Reliability Study		2nd Reliability Study	
	<u>n</u> ,	Inter-quartile range (in secs.)	<u>n</u> ,	Inter-quartile range (in secs.)
<u>1</u>	42,	-1 to 3	29,	-1 to 1
<u>2</u>	22,	-2 to 1	26,	-1 to 3
<u>3</u>	22,	-1 to 4	29,	-3 to 2
<u>4</u>	30,	-4 to 1	38,	-2 to 2
<u>5</u>	19,	-2 to 1	10,	-2 to 1
<u>6</u>	10,	-2 to 0	44,	-2 to 1
<u>7</u>	23,	-1 to 1	18,	-4 to 1
<u>8</u>	28,	-3 to 1	38,	-2 to 5
<u>9</u>	39,	-1 to 2	16,	-4 to 1
<u>10</u>	10,	-2 to 1	41,	-4 to 0
<u>11</u>	13,	-1 to 0	-	- - -

A further indication of the agreement/disagreement between observers in recording the times visitors spent viewing exhibits is provided by considering the incidence of outlying values in the various d_i distributions. In all observations, the proportion of d_i s taking values above 10 secs. or below -10 secs. was less than 10 per cent.

Observer Drifts

Some measure of observer drift in recording 'stopping' behaviour may be obtained by comparing the results from the two reliability studies. Comparison between the average interobserver percentages agreements is given in Table 2 which shows clearly that there is no significant difference between the results of the two studies at $p = .10$.

The magnitude of the bias between observers in each observation may be estimated from the statistic β , where $\beta = |a/(a + d) - \frac{1}{2}|$. In any observation, β takes a value of 0 when the likelihood of observer A recording stop is the same as observer B recording no stop and the likelihood of observer A recording no stop is the same as observer B recording stop. If the magnitude of the bias between observers has changed over time then the average values of β over observations in each study will differ (see Table 4). The mean values of β for the two reliability studies are .23 and .18. the difference between these two means is not significant at $p = .10$. ($t = .25$, d.f. = 19)

Table 4

Estimates of inter-observer bias for each observation

	1st Reliability Study	2nd Reliability Study
Observation Number	β —	β —
<u>1</u>	.5	.1
<u>2</u>	.5	.28
<u>3</u>	.07	.17
<u>4</u>	0	.1
<u>5</u>	0	.17
<u>6</u>	0	.17
<u>7</u>	.17	0
<u>8</u>	.17	.3
<u>9</u>	.1	.4
<u>10</u>	.5	.1
<u>11</u>	.5	
Mean over observations	.23	.18

To test whether the observers became more inaccurate as time elapsed in their estimates of the lengths of time visitors spent at exhibits, the average values of the magnitude of observer difference were compared between studies (see Table 5)

Table 5 (1)

Average magnitude of observer difference for each observation

Observer Number	1st Reliability Study	2nd Reliability Study
	$ \bar{d} $	$ \bar{d} $
<u>1</u>	2.0	1.0
<u>2</u>	1.6	2.1
<u>3</u>	2.7	2.7
<u>4</u>	2.6	2.3
<u>5</u>	1.4	1.8
<u>6</u>	1.6	1.5
<u>7</u>	1.8	2.4
<u>8</u>	2.5	3.2
<u>9</u>	1.8	3.5
<u>10</u>	1.1	2.1
<u>11</u>	1.1	-
Mean over observations	1.84	2.26

The means and the standard errors of the means for the two studies are $1.84 \pm .17$ and $2.26 \pm .24$. the difference between the two means is not significant at $\alpha = .10$.

- (1) The distributions of the $|d_i|$ were positively skewed. In calculating the $|\bar{d}|$ for each observation, logarithms of $|d_i|$ were taken. Table 5 shows the means of the logarithmic values of $|d_i|$ converted back to raw scores.

Discussion

The results have indicated that there is a high degree of inter-observer reliability in the use of the behavioural recording system for obtaining data relevant to measures of attracting and holding power.

Errors of commission or omission when they did occur were typically small; and upon closer inspection of the data records they were invariably the result of two distinct sorts of error. The less frequent of the two errors was an error of commission whereby an observer would incorrectly log an exhibit number code on the special keyboard. These errors are analogous to the typing errors of a typist. For example an observer would log the visitor at exhibit 13, when, in fact, he was at exhibit 14. The second sort of error was an error of omission and arose over the different interpretations by the two observers of the definition 'at an exhibit'. For visits of short duration, it transpired that on a few occasions one or other of the observers would not record the visitor 'at an exhibit' but would instead record him as 'searching'. These errors, if they occurred with any frequency would have a significant influence on computations of attracting power with the effect of decreasing the values so computed but would have an insignificant effect on computations of holding power.

Differences between the two observers in the times they recorded visitors looking at exhibits were, on average, small and randomly distributed across exhibits. That is to say, from the data collected there were no particular exhibits which attracted consistently high or low d_i s between observers. Since not all exhibits are visited by all visitors it was not possible to test the hypothesis that the different exhibits were more or less difficult to observe in a thoroughgoing fashion. (Certain exhibits were visited very infrequently by the visitors sampled

in the reliability studies.) . Furthermore, there was no indication from this study that some visitors were significantly more difficult to follow than others.

Although the results of the present study have not suggested that observer drift is a problem, it must be admitted that the design of the study was not very penetrating in this respect. Firstly, research into the phenomenon in other fields indicates that observers who know they are being assessed maintain high levels of agreement during assessment but not during covertly monitored sessions (see Hollenbeck, 1977). Unfortunately, the design of the equipment makes it impossible for a video-recording to be taken without the observer knowing about it. And, the fact that both observers claimed to be consciously less concerned and vigilant during the reliability studies than they were during the research project proper is hardly a satisfactory reason for concluding that they were, in fact, less vigilant! Secondly, only two studies were undertaken when it is advisable to make several studies throughout the course of a research project to assess the effects of observer drift over time. Like many practical research ventures, the present one was an inevitable compromise between the requirements of elegant experimental design and expediency.

Nevertheless, for the first time a serious attempt has been made to broach the problems of reliability in measuring visitor behaviour.

CHAPTER 6.

INVESTIGATING THE BEHAVIOUR OF VISITORS TO THE HALL OF HUMAN BIOLOGY.

Theoretical Background, Definitions and Notation Scheme

As pointed out in Chapter 3, the natural history of an individuals' visit to an exhibition can be thought of as a series of interactions between the visitor and the exhibits he encounters. It is by noting the exhibits a sample of visitors stops at and for how long that the indices of attracting power and holding power are calculated. Before describing the study of visitors' behaviour to the Hall of Human Biology, formal definitions will now be given of the various statistics that can be derived from recording the interactions between visitors and exhibits. These are provided in an attempt to clarify the subsequent exposition.

As Miles and Tout (1979) have suggested, it is helpful to consider the interaction between visitors and exhibits in the form of a matrix. Accordingly, suppose a population of n visitors and m exhibits, each identified by a particular value of the subscripts i and j respectively. Let the time spent by visitor i on exhibit j be t_{ij} , which in many instances is likely to be zero, and let the number of 'non-zero' interactions with exhibit j be P_j (see Figure 7). The various totals in the matrix of figures are as follows:

- T_j - total time spent by all visitors at exhibit j ,
- T_i - total time spent by visitor i at all exhibits,
- T - total time spent by all visitors at all exhibits
($\sum t_{ij}$),

P - total number of non-zero interactions with all exhibits ($\sum p_j$)

From these totals it is possible to define the 'attracting power' and 'holding power' of exhibits as follows:

P_j/n - attracting power of exhibit j ,

T_j/P_j - holding power of exhibit j ,

and P/m is the attraction of the whole exhibition and T/P in the holding power of the whole exhibition.

	Exhibit (j)										Total
	1	2	3	-	-	-	-	-	-	m	T_j
1	t_{ij}										
2											
3											
:											
:											
:											
:											
:											
:											
n											
Total	T_j										T (grand total)
	P_j										$P =$
	T_j/P_j										$T/P =$

Figure 7. Matrix for visitor/exhibit interaction (symbols explained in text.)

As described in Chapter 1, the Hall of Human Biology is partitioned into a number of smaller areas or Assemblies dealing with different topics within the overall theme with each containing different numbers of individual exhibits or displays. For ease of reference, the various assemblies and the number of exhibits in each of them are summarised below. In addition, each assembly is given an abbreviated name which will be used throughout the chapter whenever the assembly is referred to.

<u>Assembly</u>	<u>Number of Exhibits</u>	<u>Abbreviated Name</u>
A. - Living Cells	10	"Cells"
B. - Growth and Development	9	"Growth"
BE - More About Chromosomes	8	"Chromes"
C - Movement	18	"Move"
D - Controlling Your Actions	13	} "Control"
DE - More About Controlling Your Actions	3	
E - Your Life In the Balance	6	"Homstis"
F - Hormones, Messengers In the Blood	10	"Hormes"
FE - More About Sex Hormones	5	"Sexhormes"
G - Hormones and Nerves	1	"Hormnerv"
H - Learning and Memory	8	} "Mem"
HE - More About Learning and Memory	1	
I - Perception	24	"Percep"
J - How Do We Come To Understand the World About Us	25	"Cogdevel"

If individual assemblies are identified by the subscript l , then let the number of non-zero 'interactions' with Assembly l be P_l . Then the 'Attracting power' and 'Holding power' of Assemblies are

defined thus:

P_l/n - attracting power of Assembly l

T_l/P_l - holding power of Assembly l

When the attracting power of an assembly is defined in this manner, it includes visitors who might have simply walked through the assembly en route to some other destination. Accordingly, another way of considering the attracting power of an assembly is to compute the proportion of visitors visiting the assembly that actually stop at one or more of the exhibits on display. However, this measure may not give a good overall picture of an assembly's attractiveness. For example, an assembly could achieve a high overall attracting power simply on the basis of one individual exhibit which attracted all the visitors to the assembly. To take account of the relative attractiveness of the different exhibits in an assembly, a new statistic is needed.

The theoretical maximum number of visits (as opposed to visitors) to exhibits in any assembly l , will be given by $m_l n_l$ where m_l is the number of exhibits in assembly l , and n_l is the number of visitors to assembly l . This can then be related to the actual number of visits to the individual exhibits in the following manner. In any assembly, the number of visitors, n_l stopping at the j th exhibit is $n_{l,j}$ and the total number of all visits to exhibits in the assembly is given by $\sum n_{l,j}$. A new statistic, described as the Appeal of an assembly can be defined as follows:

$\sum n_{l,j} / m_l n_l$ - the Appeal of Assembly l

These statistics derived so far from Figure 7 have focused on

exhibits but by considering the rows in the matrix it is possible to derive statistics pertinent to the behaviour of individual visitors in an exhibition. For example, let P_i be the number of 'non-zero' interactions with exhibits by visitor i . Then P_i/m gives the 'Attractiveness' of the exhibition to visitor i and T_i/P_i gives what shall be styled as the 'Arrestment' of the exhibition for visitor i . Corresponding statistics can also be derived at the individual Assembly level.

Because exhibition planners are primarily concerned with mounting effective exhibits, mistakenly there has been no emphasis in the literature on the behaviour of individual visitors in the belief that little of practical value can be gained by considering individual visitors and their behaviour (Miles and Tout (1979). However, if it can be shown that the measures of 'attractiveness' and 'arrestment' vary among a population of visitors according to certain subject classifying variables (e.g. age, sex and educational attainment) then this would have implications for exhibition planning. For example, if the 'attractiveness' and 'arrestment' of an exhibition differ significantly between visitors with 'high' and 'low' educational attainment then this would suggest that it should be possible to develop different exhibitions suitable for these sub-groups of visitors.

Prior Interest in an Exhibition and its Effect on 'Attractiveness' and 'Arrestment'

In a recent paper, Griggs and Alt (In press) investigated whether visitors' assessment of how interesting they found the Museum's exhibitions once they had visited them was correlated with their prior interest in the subject matter displayed in the exhibitions, as expressed before they had visited the galleries. The authors

asked a sample of visitors entering the Museum how interested they were in seeing the various galleries and exhibitions on display. The results obtained were compared with those collected from a matched sample of visitors leaving the Museum who were asked how interesting they had found the galleries they had just seen. (This design was adopted in preference to interviewing the same sample of visitors before and after their visits to avoid the possibility of 'pre-test sensitisation'.) Alt and Griggs report that the ranking of the galleries in terms of subjects' prior interest ratings was substantially the same as the ranking of the subjects' interest measures in the exhibitions they had visited. Furthermore, this finding held among visitors coming to the Museum for the first time. Altogether, this suggests that the interests visitors bring with them ('prior interest') have a significant influence on the exhibitions they will find interesting.

The results reported by Griggs and Alt have implications for what has become known as Front-End Evaluation (see Alt, In Press). By asking visitors to a museum which of a variety of possible exhibitions they would like to see developed, a museum can more obviously develop exhibitions of known interest to its visitors. If a visitor's prior interest in an exhibition also influences his behaviour in the exhibition (such that the higher his prior interest, the higher the 'attractiveness' and 'arrestment' of particular elements of the exhibition), then the experience-value and possibly even the 'learning potential' of the exhibition is also increased.

The research reported in the remainder of this Chapter was undertaken with a number of objectives in mind. In Chapter 3, effective exhibits were identified as those that successfully

compete and attract the attention of visitors, that is, exhibits which have high attracting powers. Thus, in order to test the theory, it is necessary to compute the attracting powers of exhibits in the Hall of Human Biology. This was a fundamental objective of the present study.

It was argued in Chapter 2 that 'holding power' is a dubious theoretical concept since previous research has consistently shown that on average visitors spend only a short time viewing individual exhibits. Therefore, it was suggested that an effective exhibit is one that attracts visitors and delivers a rapid message. A further objective of the present study was to collect the time data necessary for the computation of the 'holding powers' of all exhibits thereby providing a data-base for replicating the earlier studies.

Moreover, all the studies referred to have been carried out in symmetrical, traditionally designed galleries in North American museums and art galleries. Therefore, the relevance and generalisability of their findings to a modern, didactic exhibition opened in a British museum can be examined in the light of the results of this study.

As well as collecting these basic behavioural data, the study also set out to test the hypothesis that an individual's prior interest in an exhibition is positively correlated with the 'attractiveness' and 'arrestment' of the exhibition.

Essentially, the methodology necessary to fulfil these objectives involves interviewing subjects as they enter the Hall of Human Biology (to obtain measures of prior interest) and observing their behaviour once inside the exhibition (to measure 'attracting

power', 'holding power', 'attractiveness' and 'arrestment'). Of course, as pointed out in the reference to the paper by Griggs and Alt, interviewing subjects as they enter an exhibition about their interests in the various topics on display runs the danger of pre-test sensitization. In the present case, the fact that a subject expresses a particular interest in a certain section of the exhibition might compel him to behave fairly consistently once he is in the exhibition, thus engendering some circularity in the data. Theoretically, it would have been possible to observe the behaviour of subjects in a suitably matched control group who had not been interviewed and subsequently to have compared their behaviour with that of the experimental subjects. This method was not adopted for the following reasons. With such a heterogeneous population of visitors to be covered - casual museum visitors over the age of eleven years - quite large samples would have been required to match two groups on 'prior interest'. This is true also if visitors were randomly allocated to groups; even then, there is no guarantee that the two groups would have been matched on 'prior interest' unless large samples were used. The time available for the study simply precluded the possibility of observing two large groups. At the time, it was also felt that the possible effects of pre-test sensitization were small, although there is, of course, no logical reason why this should be so. In brief, then, the approach adopted involved interviewing subjects before they entered the exhibition and subsequently, using the CCTV system described in Chapter 4, unobtrusively observing their behaviour in the exhibition.

In summary, the major objectives of the study were:-

1. to measure the 'attracting powers' and 'holding powers' of

the various exhibits and assemblies;

2. to measure the 'attractiveness' and 'arrestment' of the exhibition
3. to test the hypotheses that subjects' prior interest in the exhibition was positively correlated with its 'attractiveness' and its 'arrestment'.

Method

Subjects

Three hundred and seventy eight visitors to the Hall of Human Biology during November and December 1979, and January, February and March, 1980 were selected systematically as they were entering the exhibition, using the sampling method described in Appendix I. Essentially, the method requires that only visitors entering the exhibition at pre-determined times are approached, thus taking the choice of subjects out of the hands of the interviewers. The following visitors were excluded from the study:-

- (i) non-English visitors;
- (ii) visitors under the age of eleven years;
- (iii) schoolchildren in organised school parties (members of sixth form or student groups, however, were not excluded if entering the exhibition in small groups of fewer than five people);
- (iv) teachers accompanying organised parties of school-children;
- (v) visitors entering the exhibition for the second

time in one day;

- (vi) visitors intending to leave the Museum within an hour (since it was argued that visitors who knew that they had to leave the Museum within a short period of time would not have the time to look at the exhibition even if they wanted to.)
- (vii) visitors entering the exhibition expressly to find a member of their group rather than to look at the exhibition.

The breakdown of visitors approached during the period was as follows:-

Interviewed	148
<hr/>	
Refusals (did not want to be interviewed)	10
Non-English speaking	102
Teachers accompanying school parties	2
Visitors re-entering the exhibition on the day of interview	6
Visitors with less than one hour to spend in the Museum	110
	<hr/>
Total approached:	378
	<hr/>

The profile of the sample of 148 subjects interviewed is given in Table 6 and alongside the profile of the sample of visitors interviewed in the 1979 Visitor Survey (Appendix I) is also presented. It can be seen that there is a good degree of similarity between the two samples, something which affords confidence in the sampling method used, at least in its consistency. These data also indicate that visitors to the Hall of Human Biology are representative of all visitors to the Museum.

Table 6

Comparison between Sample of Subjects and Visitors interviewed in 1979 Survey

	Hall of Human Biology Study 1	1979 Visitor Survey
Subject Characteristics	n = 148 %	948 %
<u>Sex</u>	Male 55 Female 45	53 47
<u>Age</u>	11 - 16 years 5 17 - 20 years 17 21 - 24 years 16 25 - 34 years 25 35 - 44 years 19 45 - 54 years 11 55 - 64 years 4 65+ years 2	4 15 19 29 18 8 5 2
<u>Group Composition</u>	Alone 36 Organised Party 7 Friends 16 Family only 39 Family/Friends 2 ----- Without children 68 With children 32	27 2 23 44 4 ----- 62 38
<u>Distance of Residence from Museum</u>	Overseas 40 0 - 10 miles 14 10 - 20 miles 11 21 - 30 miles 6 31 - 40 miles 4 41 - 50 miles 1 51 - 100 miles 5 100 - 200 miles 5 Over 200 miles 14	42 14 9 5 2 3 7 10 8

1. Additional information collected about the 148 subjects was as follows:

- 80 claimed to have a science qualification;
- the mean time available in the Museum was 2.43 hours (S.D. 0.93 hours);
- 133 were visiting the Hall of Human Biology for the first time;
- 60 intended to visit (or had already visited) another museum on the day they were interviewed;
- 48 had not visited any galleries in the Museum before visiting the Hall of Human Biology on the day they were interviewed;
- 124 wished to visit other galleries before they left the Museum.

Measures Taken

The one hundred and forty eight subjects who were interviewed were asked to state which of five statements best described their interest in seeing the various topics on display in the Human Biology exhibition to yield a measure of their prior interest in the exhibition.

Interest in an exhibition can be conceived on a scale ranging from extreme disinterest to extreme interest, and an appropriate scaling method can then be used to measure the extent of the interest. However, in all museum studies undertaken at the Natural History Museum there has been a marked reluctance on the part of visitors to be critical of Museum displays when questioned (see Alt, In Press). This response bias resulted in a skewed distribution of responses to a bipolar scale of interest with very few visitors stating they were disinterested in an exhibition they were about to see or indeed had seen. In the study reported by Griggs and Alt (In Press) subjects were asked to state how interested they were in seeing the Museum's exhibitions and galleries by choosing one of the following five statements:

<u>Statement</u>	<u>Scale Value</u>
Not particularly interested	1
Not sure	2
Slightly interested	3
Quite interested	4
Extremely interested	5

The results they obtained are summarised in Table 7.

Table 7

Expressed Interest in the Exhibitions/Galleries (From Griggs and Alt, In Press)

$n = 301$	Not particularly Interested	Not sure	Slightly Interested	Quite Interested	Extremely Interested	Mean Rating
	%	%	%	%	%	
1. Whale Hall	4	3	15	39	39	4.1
2. Dinosaurs	6	1	18	42	33	4.0
3. Mammals	5	3	12	51	29	4.0
4. Origin of Species	6	4	15	34	42	4.0
5. Man's Place in Evolution	7	5	16	38	34	3.9
6. Human Biology	13	5	17	36	30	3.6
7. Fossil Galleries	22	6	25	29	18	3.2
8. Birds	21	6	26	33	14	3.1
9. Nature Stored, Nature Studied	15	13	27	34	12	3.1
10. Fish and Reptiles	22	6	32	29	11	3.0
11. Minerals and Meteorites	31	7	23	23	16	2.9
12. Insects	33	6	22	29	11	2.8
13. Introducing Ecology	28	18	21	22	10	2.7
14. Botany	37	9	26	21	6	2.5

Of course, the nature of the scale used in this study will have had an influence on the results obtained. However, the scale was designed in an attempt to provide a reasonable discrimination between subjects in terms of their prior interest in the exhibitions on display. As such, the authors believed it had heuristic value even though it is a rather inelegant example of scaling methods. It can be seen from Table 7 that a fair spread of average values over the various galleries was obtained. In earlier pilot attempts to discriminate between galleries using bipolar 5-point or 7-point scales of interest, subjects tended to cluster around the positive end of the scale. This sort of positive response bias among visitors asked to evaluate exhibitions had been found to be so strong that when 'very' interesting was used as the upper scale value (rather than 'extremely'), the majority of visitors tended to opt for the 'very' position rather than the 'quite' (interesting) position preceding it. Thus, the use of 'extremely' interesting was designed specifically in an attempt to inhibit subjects from choosing the upper limit of the scale.

The scale used by Griggs and Alt was the one used in the present study to measure subjects' prior interest in the Human Biology exhibition. Subjects were asked to rate how interested they were in seeing each of the assemblies; and an individual's prior interest in the exhibition was defined, accordingly, as the mean value of his ratings of all assemblies.

During the interview, certain characteristics of the subjects were also collected (e.g. age, sex and science qualifications), whereupon subjects entered the exhibition and their behaviour during their visits was observed unobtrusively to yield measures of 'Attracting power', 'Holding power', 'Appeal', 'Attractiveness' and 'Arrestment' as described and defined earlier in the chapter.

The precise procedure adopted for selecting and interviewing subjects is now described.

Procedure

Selected subjects were approached as they were about to enter the Hall of Human Biology. The interviewer introduced herself to the visitor and asked if he/she would be willing to answer some questions. The exact form of introduction was cast as follows:

"Excuse me, my name is and I work here at the Natural History Museum. We are carrying out a study of visitors to the Hall of Human Biology and I should be grateful if you would help me by answering a few questions?"

Upon obtaining agreement, the interviewer established whether the visitor was in exclusion categories (i) - (vii), and if so the interview was abandoned. If not, he or she was taken to an interview room adjacent to the exhibition. (The interview room was the room containing the real time observational system described in Chapter 4. However, the apparatus was shielded from subjects by screens.) He or she was then invited to sit down in a comfortable chair and the interviewer sat opposite, across a coffee table. The interviewer then asked the following questions:

"Have you visited any other galleries in the Museum today before coming to the Hall of Human Biology?"

IF YES: "How many?"

"Are there any other galleries you would certainly like to

see after your visit to the Hall of Human Biology if you have any time left on your visit today?"

"Have you ever visited the Hall of Human Biology before your visit to the Museum today?"

IF YES: "How many times have you visited the Hall of Human Biology before today?"

In order to obtain the key 'Prior Interest' measure subjects were then handed a set of cards on which were written descriptions of the topics covered in each of the assemblies of the exhibition,¹ and were asked the following:

"The Hall of Human Biology is a large exhibition covering a number of different topics. Each topic is arranged in a separate section of the exhibition. On these cards are written brief descriptions of the topics covered in each of the sections in the exhibition. I'd like you to go through the cards and tell me how interested you are in seeing each of these sections on your visit today?"

At the same time, subjects were shown a card on which were printed the statements "Not particularly interested", "Not sure", "Slightly interested", "Quite interested", "Extremely interested", and they were invited to indicate which statement best described their interest in seeing each of the sections.

1. The descriptions were written by the team of scientists responsible for the content of the exhibition. They referred only to the topics and no mention was made of the design of the exhibitions. The descriptions are given in Appendix II.

Next, they were asked to do the following task:

"Now, I'd like you to pick out the card which describes the section you think you would be most interested in seeing on your visit today."

When the subject had selected a card, he or she was asked to proceed as follows until a complete ranking of assemblies was obtained.

"Of the remainder, please pick out the card which describes the section you think you would be most interested in seeing on your visit today."

The cards were shuffled between subjects to ensure randomisation. The interview was concluded with a set of questions to measure certain characteristics of each subject (e.g. age and science qualifications). These characteristics were similar to those collected on the annual visitor surveys. (see Appendix I).

Upon completion of the interview, subjects were escorted to the entrance of the exhibition by the interviewer.

As they entered the exhibition and for the duration of their time in it, they were observed by a research assistant using the apparatus described in Chapter 4. Thus, the assistant observed subjects as they wandered through the exhibition on closed-circuit television, switching between cameras to keep track of them while at the same time logging their behaviour on the keyboard specially constructed for the purpose. The exhibits they stopped at and for how long, were recorded on the real-time system.

The data from the interviews together with the observational data for each subject were placed on to code sheets and later transferred to punch cards for analysis.

Results

Introduction

In presenting the results of this study a hierarchical plan will be followed, starting at measures relating to the exhibition as a whole, then moving to the level of assemblies (groups of exhibits) and finally to the level of measures concerning individual exhibits. Table 8 shows this presentation.

Table 8

Plan of Data Presentation

Exhibition Measures	Prior interest Length of visit Attractiveness Arrestment
Assembly Measures	Prior interest Time spent at assemblies Total interactions:- <ul style="list-style-type: none">- proportion of subjects visiting- proportion of subjects stopping- appeal Holding power
Exhibit Measures	Attracting power Holding power

Exhibition Measures

Prior Interest

Figure 8 shows the distribution of prior interest in the exhibition among the sample of 148 subjects. The mean and standard deviation of the distribution are 3.6 and .64 respectively, and suggest the response bias mentioned earlier by showing that typically subjects have a tendency to express a strong interest in the topics on display in the exhibition. Nevertheless, the bias was not nearly as marked as that found using bi-polar scales of interest and the results indicate that the scale adopted had a reasonable heuristic value in discriminating between subjects.

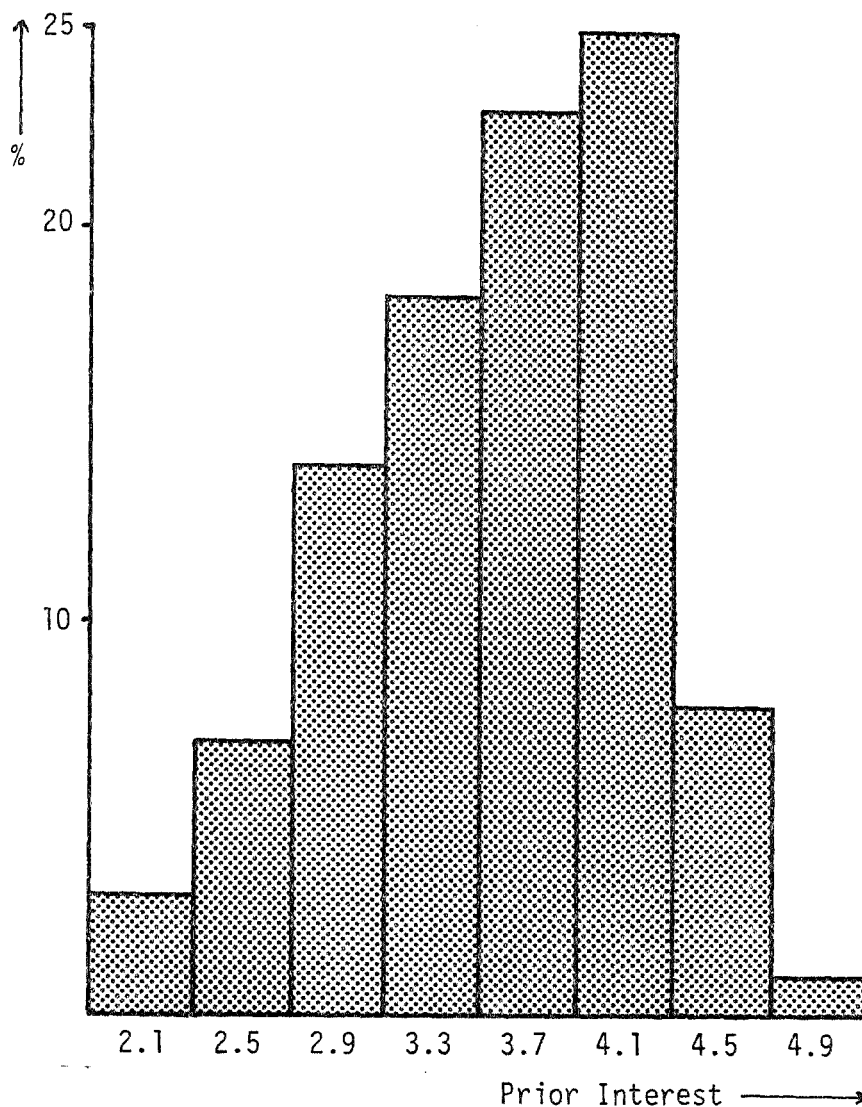


Fig. 8. Prior interest in the Hall of Human Biology.

Duration of visit to the Hall of Human Biology

The distribution of visiting times among the sample is shown in Figure 9. In view of the skewness of the distribution a variety of summary statistics is given in Table 9. Table 9 shows the total time spent in the exhibition broken down into the average time spent looking at exhibits and the average time spent wandering through the exhibition (styled as 'searching' in the Table).

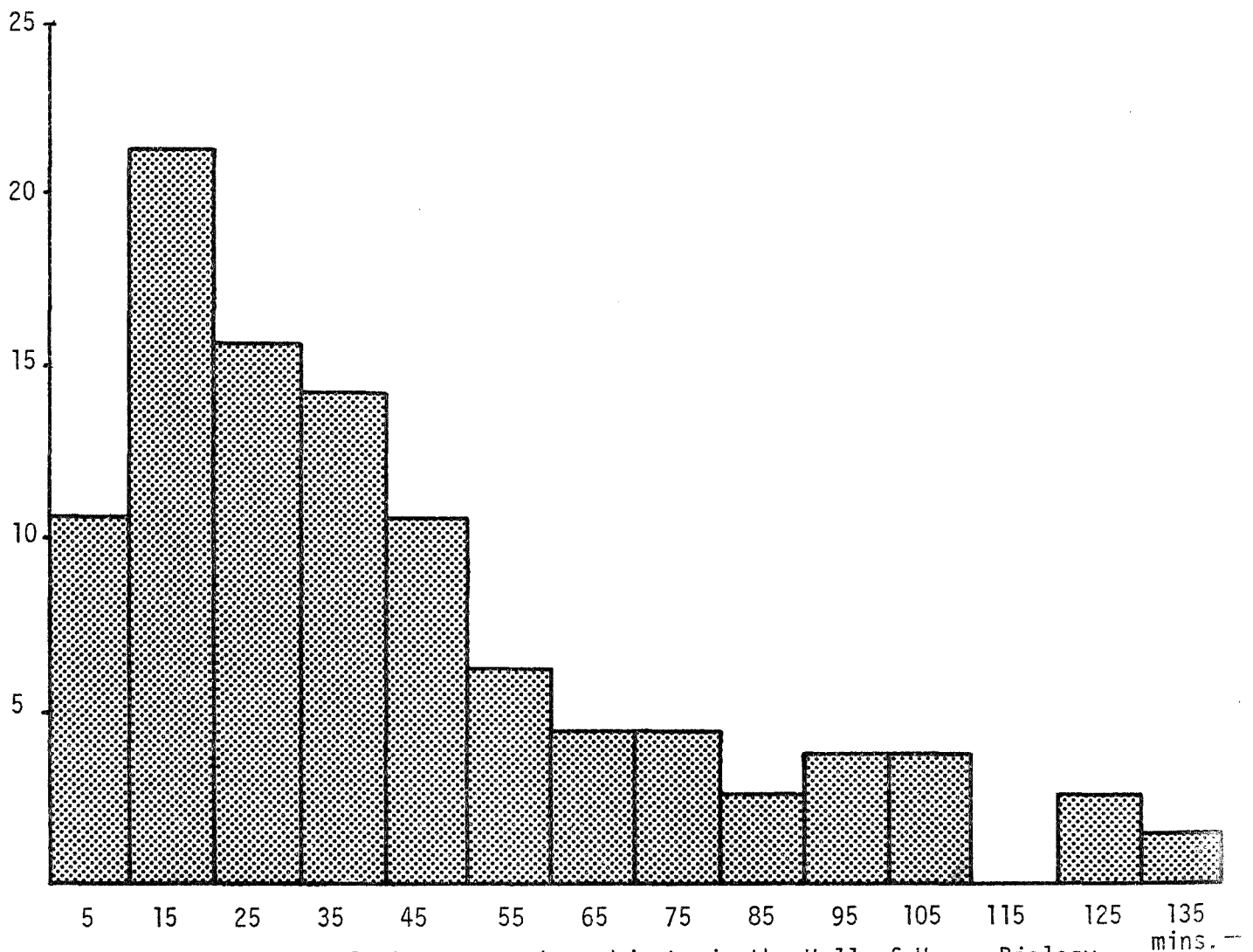


Fig.9 Distribution of times spent by subjects in the Hall of Human Biology.

Table 9

Time spent in the Hall of Human Biology (in minutes)

	Mean	S.D.	Median	Inter Quartile Range
Total time spent in exhibition	38.9	30.7	30.5	16.7 - 52.1
Time spent at exhibits	29.1	27.9	20.8	10.3 - 40.8
Time spent searching	9.1	4.7	8.9	6.0 - 11.9

The average amount of time available in the Museum to subjects was 2.43 hours (see footnote to Table 6) On average approximately one quarter of that time was spent in the Hall of Human Biology.

Attractiveness of the Hall of Human Biology

The distribution of the attractiveness of the exhibition among the sample is shown in Figure 10. Attractiveness varies in a range from 0 to 0.9 with a mean and S.D. of 0.36 and 0.18, respectively. Though slightly skewed, the distribution is reasonably well described by the mean and standard deviation. Thus, on the average visitors look at one third of the exhibits on display in the Hall of Human Biology.

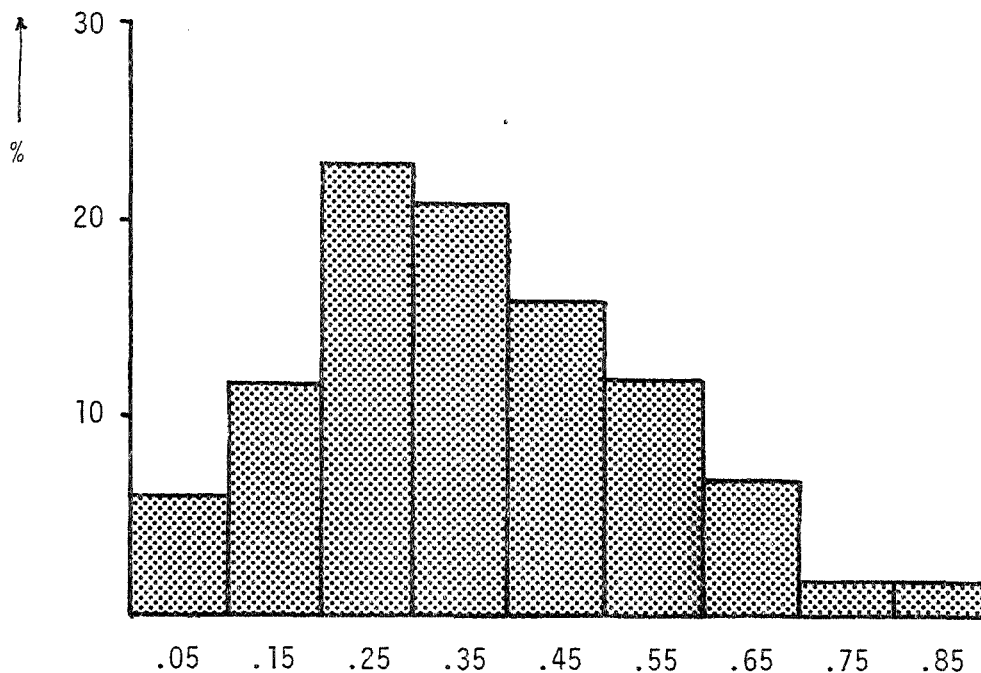


Fig. 10. Distribution of Attractiveness of the Hall of Human Biology.

Arrestment of the Hall of Human Biology

The distributions of individual subjects' exhibit viewing times were positively skewed¹ and, therefore, each individual's viewing times were converted to natural logs before computing his 'arrestment' value. The distribution of the log values of arrestment are shown in Figure 11. When converted back to raw scores, the range of arrestment was 0 to 65 seconds, with a mean of 18 seconds. The ranges of the means ± 1 S.D. and means ± 2 S.D. were respectively, 11 to 30 seconds and 6 to 51 seconds.

-
1. Typically, the distribution of an individual subject's viewing times included a few high values (that is, a few exhibits he or she viewed for much longer than the rest of the exhibits stopped at). Therefore, if arrestment were calculated on raw scores, the values would be unduly influenced by the small number of relatively extreme scores, resulting in a high and unrepresentative value of arrestment (see Inset to Fig. 11 for distribution of arrestment on raw scores.)

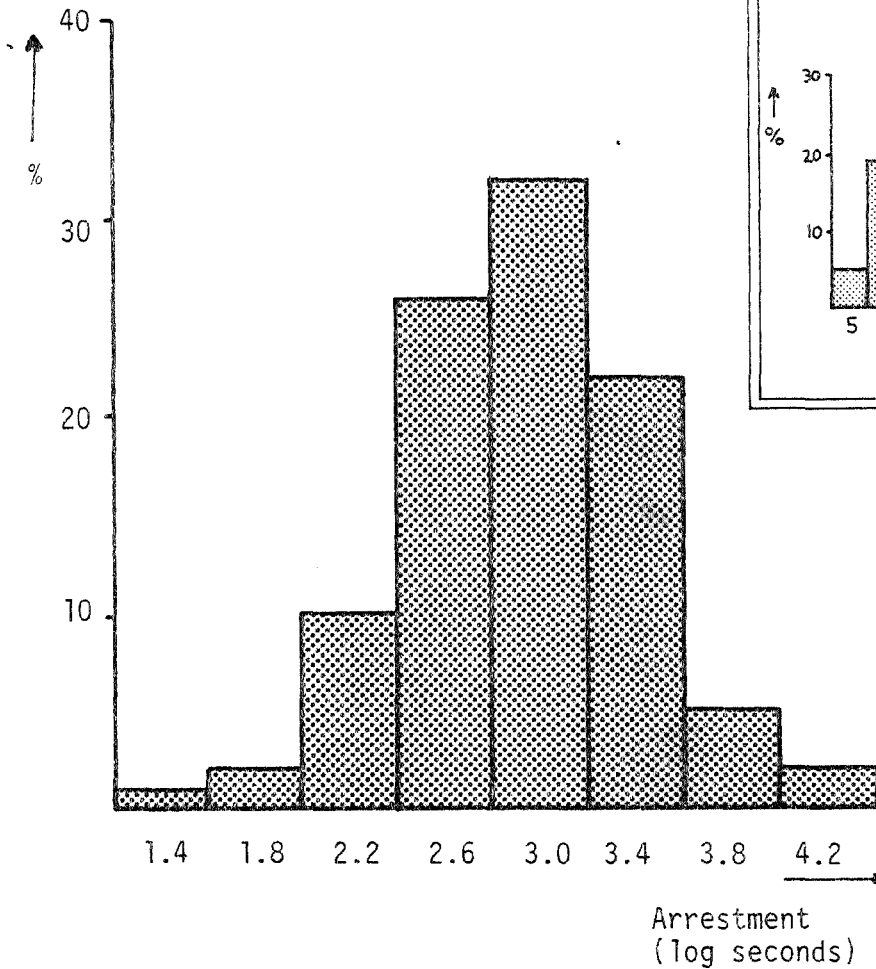


Fig. 11. Distribution of Arrestment in the Hall of Human Biology (in log seconds) (Inset shows distribution of arrestment in raw scores)

Assembly Measures

Prior Interest

The data on prior interest also indicated the relative interest subjects had in the various topics on display in the different assemblies. In this instance, averages were calculated on the scale ratings (5-point scale) and on the ranks (1 to 12) given to each assembly. These results, together with the correlations (Pearson's r) between rating and ranking scores are presented in Table 10.

Of the topics covered in the different assemblies, the ones which held most interest for subjects were those dealing with psychological topics rather than the strictly biological. These were:-

- assemblies covering the topics of learning, memory and recall (H and HE);
- aspects of cortical control (D and DE);
- perception (I); and
- cognitive development (J).

Inspections of the correlations between ranking and rating scores also reflect the response bias obtained by the rating measure. Specifically, the interest measured by rating is higher than that measured by ranking, resulting in relatively low values of the various values of r . Also the magnitude of the standard deviations suggest clear differences of opinion among subjects.

Table 10

Average measures of prior interest for each assembly

Assembly	Rating (1 - 5)		Ranking (1 - 12)		Correlation Rating:Ranking r
	Mean	S.D.	Mean	S.D.	
'Mem' (H & HE)	4.12,	0.94,	8.56,	2.85,	0.53
'Control' (D & DE)	3.93,	0.90	7.46,	2.76,	0.30
'Percep' (I)	3.83,	1.07,	7.69,	3.06,	0.56
'Cogdevel' (J)	3.71,	1.21	7.39,	3.41,	0.50
'Hormnerves' (G)	3.12,	1.40,	5.25,	3.31,	0.60
'Hormes' (F)	3.26,	1.33,	5.68,	3.08,	0.55
'Move' (C)	3.27,	1.22,	5.79,	3.38,	0.47
'Sexhormes' (FE)	3.27,	1.28,	5.27,	3.10,	0.60
'Homstis' (E)	3.23,	1.11,	5.79,	3.13,	0.49
'Cells' (A)	3.33,	1.23,	5.91,	2.73,	0.45
'Growth' (B)	3.44,	1.32,	6.41,	3.49,	0.68
'Chromes' (BE)	3.48,	1.29,	6.46,	3.52,	0.55

Interactions with assemblies

In the first section of this chapter, three different measures of the attracting power of an assembly were defined, namely (1) the proportion of visitors to an assembly, (2) the proportion of visitors stopping at one or more exhibits in an assembly, and, (3) the appeal of an assembly which was a new statistic defined to take account of the relative attracting powers of the different exhibits in an assembly. Table 11 shows the three different measures.

Table 11

Attracting powers of assemblies in the Hall of Human Biology

Assembly	(1) Proportion visiting	(2) Proportion of visitors stopping at one or more exhibits	(3) Appeal
'Cells' (A)	1.00	.92	.46
'Growth' (B)	.99	.92	.47
'Chromes' (BE)	.69	.54	.18
'Move' (C)	.93	.91	.32
'Control' (D & DE)	.80	.94	.43
'Homstis' (E)	.92	.72	.26
'Hormes' (F)	.89	.94	.42
'Sexhormes' (FE)	.57	.79	.41
'Hormnerv' (G)	.42	.63	.63
'Mem' (H & HE)	.86	.75	.42
'Percep' (I)	.68	.86	.38
'Cogdevel' (J)	.74	.93	.35

The appeal of an assembly is the same statistic as the attracting power of an assembly when the assembly contains only one display. This is the case with assembly G ('Hormnerv'). Assemblies C, I and J ('Move', 'Percep' and 'Cogdevel', respectively) contain 18, 24 and 25 exhibits, which are relatively high numbers in comparison to other assemblies. Appeal seems to some extent to be a function of the number of exhibits in an assembly since the assembly with the fewest exhibits, G ('Hormnerv'), has the highest appeal and the appeals of the assemblies with the most exhibits, I and J ('Percep' and 'Cogdevel'), tend to fall. However, the relationship between the number of exhibits in an assembly and its appeal is not straightforward. It is, therefore, advisable to view appeal with some caution as it may give a distorted picture of the attractiveness of assemblies with small and large numbers of exhibits. However, taking all measures together, it is clear that BE and E, ('Chromes' and 'Homstis') are the least attractive assemblies, but the question as to which are more attractive is complicated. A and B ('Cells' and 'Growth') emerge as attractive on all measures but they are the first two assemblies visited by the majority of visitors. D and DE ('Control') appear as reasonably attractive as does F ('Hormes'); and perhaps C ('Move') is less attractive than I and J ('Percep' and 'Cogdevel') in view of the fact that the latter two assemblies are to be found at the end of the exhibition.

Holding Power

The holding power of an assembly was defined as T_l/P_l , where T_l was the time spent by subjects in assembly l and P_l was the number of 'non-zero' interactions with exhibits in assembly l . Table 12 shows the time spent looking at exhibits (holding power in the strict sense) and the time spent searching or walking between

exhibits. These summary statistics indicate skewness in the distributions of holding power and, therefore, the median may be taken as a better estimate of central tendency than the mean. Clearly, assemblies I and J ('Percep' and 'Cogdevel') have the highest holding powers; and these are the two assemblies containing the largest number of exhibits. The figures for searching time generally show less variation and are closely related to the number of exhibits in each assembly, reflecting the time it takes to circumnavigate the floor area.

Before proceeding to an analysis of the interactions between subjects and individual exhibits, it is worth remarking that the data just presented at the assembly level have indicated that attempts to draw conclusions about the relative attracting powers and holding powers of assemblies have been largely unfruitful. The confounding variable in these attempts has been the different numbers of exhibits on display in each assembly. The data have indicated that there are differences between assemblies, particularly in terms of the various measures of attracting power, but the effects have been confounded by the variable amounts of information on display in each assembly. Of course, there is no way of controlling for this variable and the assembly is probably, therefore, a futile level of analysis. Logically, it is better to go down to the exhibit level; and psychologically this also makes better sense since, after all, individual exhibits are the 'objects' that visitors stop and look at. Furthermore, the theory of exhibit effectiveness outlined in Chapter 3 is concerned with the perception of individual exhibits in competition with other exhibits.

Table 12

Holding Powers of Assemblies in the Hall of Human Biology

Assembly	Time at Exhibits (mins.)				Time searching (mins.)			
	n	Mean	S.D.	Median	n	Mean	S.D.	Median
A	136	2.5	2.3	1.6	148	1.1	.54	1.1
B	136	3.4	3.9	1.8	147	.68	.57	.58
BE	55	4.4	5.9	1.8	102	.65	.60	.42
C	126	2.5	2.5	1.8	138	.98	.65	.92
D	112	3.2	3.0	2.6	119	.97	.98	.75
DE	63	1.4	1.4	.92	66	.21	.14	.22
E	98	1.2	1.2	.58	136	.42	.37	.30
F	123	3.1	3.0	2.3	136	1.3	1.2	.92
FE	66	1.7	1.9	.92	84	.39	.23	.42
G	39	.63	.90	.22	61	.07	.06	.05
H	95	3.6	3.6	2.1	112	1.04	.74	.75
HE	68	1.7	1.5	1.3	127	.27	.71	.13
I	87	5.5	4.3	4.6	101	1.5	.79	1.4
J	102	10.1	11.1	4.6	110	1.7	1.0	1.4

Exhibit Measures

Attracting Powers

The attracting power of an exhibit was defined as P_j/n where P_j was the number of non-zero interactions with exhibit j by a sample of n visitors. The distribution of attracting powers calculated in this manner is shown in Figure 12. The distribution, though slightly skewed is reasonably well described by the mean and standard deviation. These were .31 and .15 respectively.

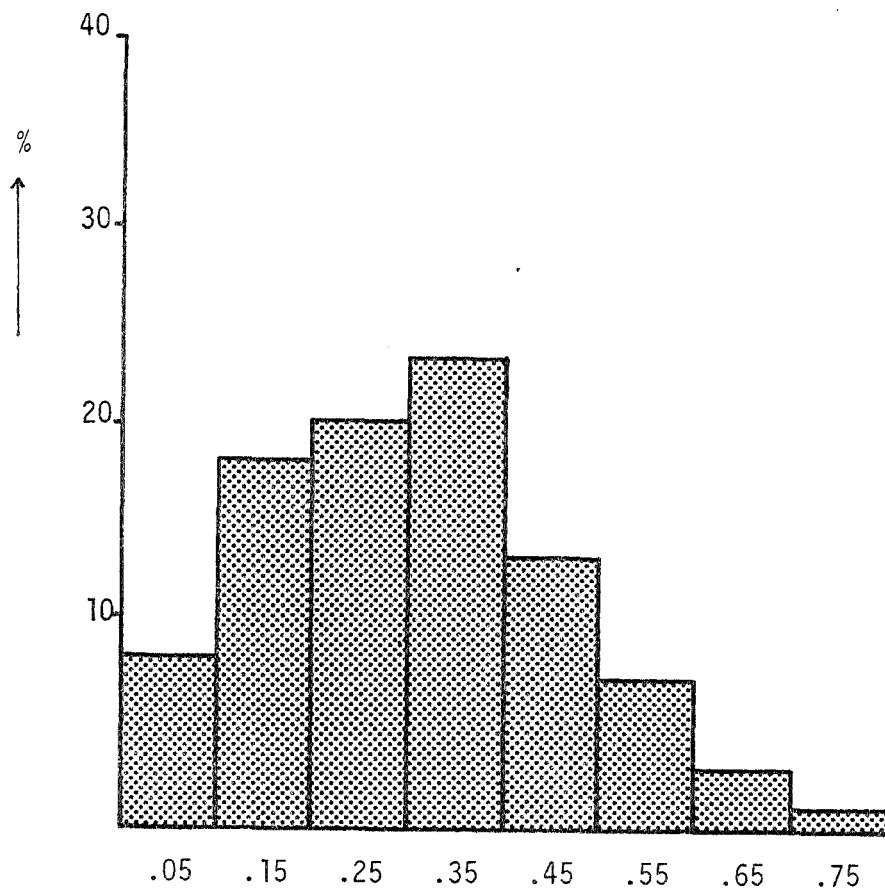


Figure 12. Distribution of exhibit attracting powers in the Hall of Human Biology.

Holding Powers

Typically, the distributions of times spent by subjects viewing individual exhibits were skewed. This had been found in other observational studies carried out in the Natural History Museum (Alt, 1979). Therefore, median values were taken as a better measure of holding power than the arithmetic mean defined earlier in the chapter. (An alternative would have been to have taken the mean value of the logarithmically transformed raw scores; this was done in calculating arrestment.) The distribution of the holding powers (medians) of exhibits is shown in Figure 13. As can be seen, the distribution is roughly log-normal, hence on a log scale, the distribution is fairly well described by the mean and standard deviation. The ranges of the mean \pm 1 S.D. and the mean \pm 2 S.D.s are 9, 19, 42 and 4, 19 91, respectively when converted back to raw scores (medians in seconds).

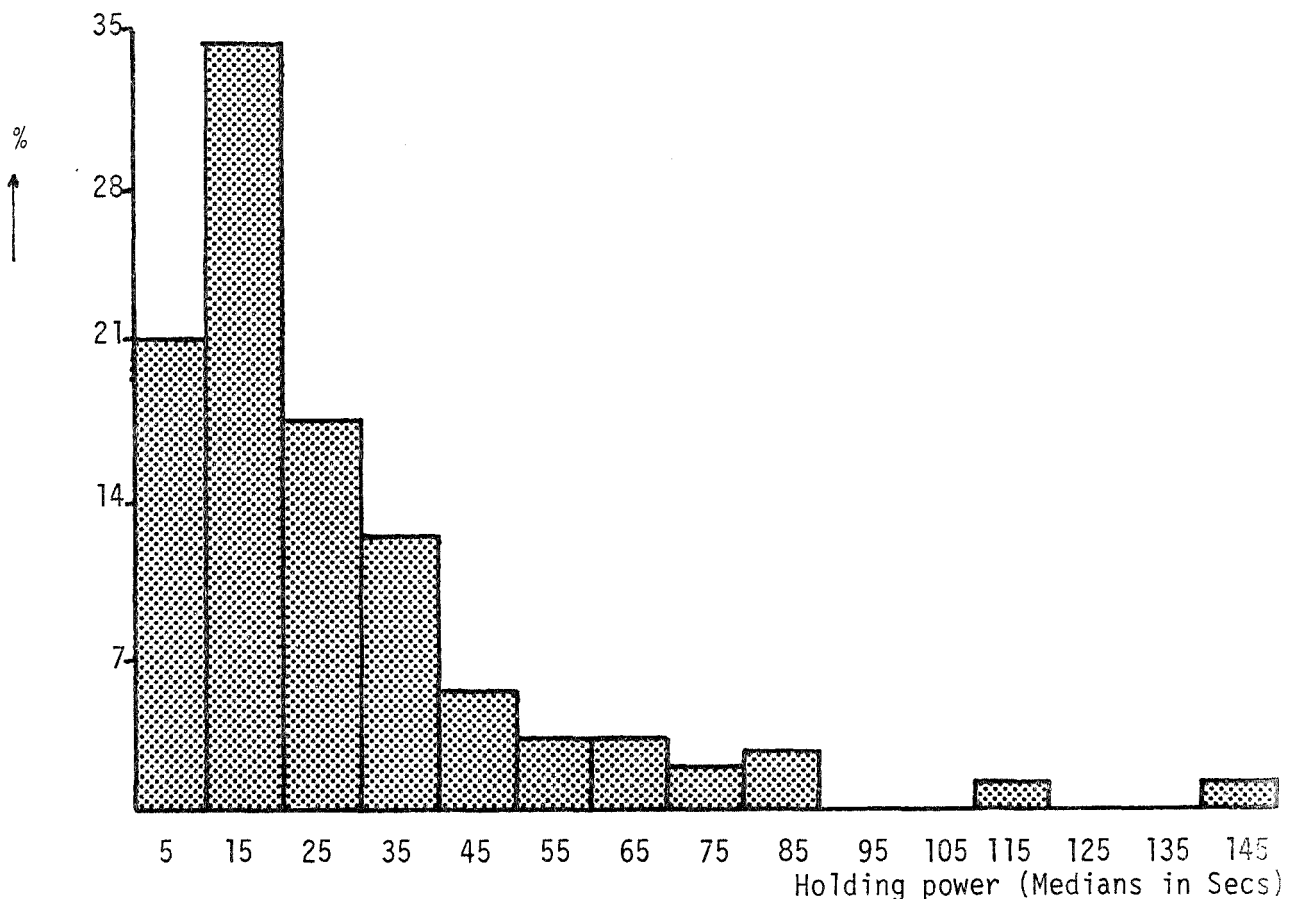


Figure 13. Distributions of exhibit holding powers in the Hall of Human Biology.

Table 13

Summary of behavioural data: . Assemblies and Exhibits

Level of Analysis	Holding Power (medians)	Attracting Power		
		Proportion visiting	Proportion of visitors stopping	Appeal
Exhibition	20.8 mins	1.00	1.0	na
<u>Assemblies</u>				
A ('Cells')	1.6 mins	1.00	0.92	0.46
B ('Growth')	1.8 mins	0.99	0.92	0.47
BE ('Chromes')	1.8 mins	0.69	0.54	0.18
C ('Move')	1.8 mins	0.93	0.91	0.32
D]	2.6 mins	0.80	0.94	0.43
DE] - ('Control')	.92 mins			
E ('Homstis')	.58 mins	0.92	0.72	0.26
F ('Hormes')	2.3 mins	0.89	0.94	0.42
FE ('Sexhormes')	.92 mins	0.57	0.79	0.41
G ('Hormnerv')	.22 mins	0.42	0.63	0.63
H]	2.1 mins	0.86	0.75	0.42
HE] - ('Mem')	1.3 mins			
I ('Percep')	4.6 mins	0.68	0.86	0.38
J ('Cogdevel')	4.6 mins	0.74	0.93	0.35
Exhibits	Average Holding Power (medians)	Average proportion of exhibits stopped at		
	19 secs	0.31		

Table 13 summarises the behavioural data about assemblies and exhibits collected from the sample of 148 subjects observed during its visit

to the Hall of Human Biology. On average visitors spend approximately twenty one minutes looking at the exhibits on display in the whole exhibition. The assemblies in which visitors typically spent most of their time looking at exhibits were I ('Percep') and J ('Cogdevel') but these are the assemblies containing the most exhibits (approximately a third of all exhibits are found in these two assemblies) and on average approximately twenty per cent of visitors' time is spent looking at exhibits in them. The typical visitor looked at fewer than a third of the exhibits on display and spent less than twenty seconds viewing individual exhibits. Of course, there were considerable individual differences and some visitors did spend considerably longer looking at certain exhibits but on the basis of the behavioural data alone, the conclusion must be that designers of ambitious didactic exhibits like the Hall of Human Biology profoundly overestimate the efforts visitors are prepared to invest in looking at the exhibits on display.

Individual differences in Attractiveness and Arrestment of the Hall of Human Biology.

Earlier in the chapter, the attractiveness of the exhibition to a visitor was defined as P_i/m , where P_i is the number of 'non-zero' interactions with m exhibits by visitor i ; and the arrestment of the exhibition was defined as T_i/P_i , where T_i is the total time spent in the exhibition by visitor i . The average attractiveness of the Hall of Human Biology was found to be 0.36 and the average arrestment was 18 seconds. (c.f. the corresponding statistics for the average attracting power and holding power of exhibits - 0.31 and 18 seconds, respectively). Despite these relatively low values, some encouragement would be given to exhibition planners if the attractiveness and arrestment of the exhibition proved to be greater among certain groups of visitors since an implication of such a finding would be that different exhibitions could be

developed for these groups.

The visitor surveys conducted by the author at the Natural History Museum (see Appendix I) had suggested that segmentation of the population in terms of demographic characteristics results in no obvious different preferences or opinions. The educated and the uneducated, the old and young alike all appeared to like the same exhibitions although how much was 'factual' and how much due to instrument insensitivity was not known. However, a more recent survey by Griggs and Alt (in Press) already referred to, has shown that visitors' prior interest in the topics on display influenced their preference of the galleries they subsequently visited, to the extent that of the various galleries they visited, visitors apparently preferred the galleries dealing with topics in which they had a prior interest.

The present study was designed partly to build upon this finding by investigating the hypothesis that the prior interest an individual has in the exhibition is positively correlated with the attractiveness and arrestment of the exhibition. In this section, the relationship between attractiveness and prior interest will be considered first, followed by the relationship between arrestment and prior interest.

In addition to an individuals prior interest, defined earlier as the mean value of his ratings of prior interest in all the different assemblies, a number of other individual difference variables were also correlated with attractiveness and arrestment. The variables and their definitions were:-

- sex (male or female);
- age (11 - 16 years, 17 - 20 years, 21 - 24 years, 25 - 34 years, 35 - 44 years, 45 - 54 years, 55 - 64 years and 65+ years);
- qualifications in science (qualifications or no qualifications);
- visit status (visited the Museum alone or with others).

The correlations between attractiveness and the various subject variables are given in Table 14. A significant positive correlation was obtained between attractiveness and prior interest ($r = 0.25$, $P < 0.01$), suggesting that the attractiveness of the Hall of Human Biology is to some extent a function of the subject matter displayed though only $0.25^2 = 6.25\%$ of the common variation is predictable from the correlation between arrestment and prior interest. There was also a significant negative correlation between arrestment and visit status ($r = -0.18$, $P < .05$) suggesting that accompanied visitors are attracted to fewer exhibits than solitary visitors. This finding echoes the work of Lakota (1975) who found that accompanied visitors spent a significant amount of their time in an exhibition interacting with one another rather than the exhibits on display, an opportunity clearly not available to the solitary visitor.

Table 14

Correlations between attractiveness and individual difference variables

	Sex	Age	Qualifs. in Science	Visit Status	Prior Interest
Attractiveness	0.01	-0.13	-0.08	-.18*	0.25**

** significant at 1% level

* significant at 5% level

Next, the correlations between arrestment and the various subject variables were computed. No significant correlations were obtained. The full set of correlations is given in Table 15.

Table 15

Correlations between arrestment and individual difference variables.

	Sex	Age	Qualifs. in Science	Visit Status	Prior Interest
Arrestment	.04	.03	.14	.06	.16

These results have provided some evidence that the prior interest a visitor has in the subject matter displayed has a significant if not very substantive influence on his or her behaviour in the exhibition. Specifically, the greater the prior interest an individual has, the greater the attractiveness of the exhibition. However, the prediction of attractiveness from prior interest is weak which suggests that the dynamics of the visiting process, e.g. the differential impulsivity of attraction to particular exhibits, needs to be taken into account in a complete description of what makes certain exhibits rather than others attractive. This point will be taken up later in the Discussion section and is the major preoccupation of the next chapter. The fact that prior interest had no influence on arrestment is of significant theoretical interest. It was mentioned in Chapter 2 that the holding power of an exhibit was a dubious theoretical construct given that all the studies reported in the literature indicated that the average holding power of exhibits was less than sixty seconds. In the present study, these apparently immutable findings have been replicated.

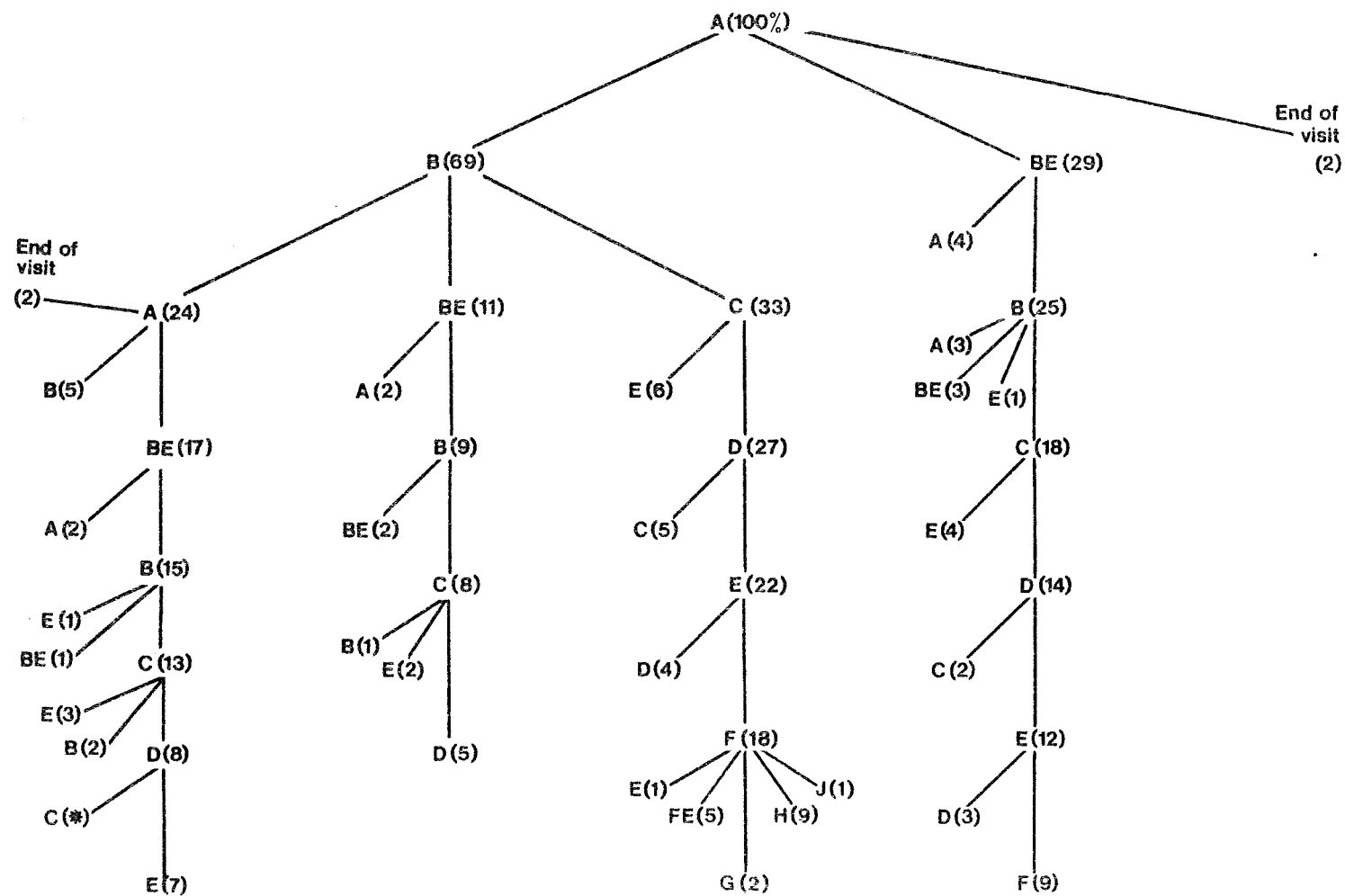
Ninety five per cent of visitors were found to spend on average between 6 and 51 seconds looking at an exhibit in the Hall of Human Biology and this was virtually irrespective of their expressed interest in the subject matter displayed in the exhibits. It seems that there is a critical time beyond which visitors are simply not prepared to stay at any one exhibit.

Before discussing the relevance of these results to the proposed theory of exhibit effectiveness and to exhibition planning in general, the remainder of this chapter is devoted to examining the routes taken by visitors through the Hall of Human Biology. As well as being important in its own right (since the exhibition was developed on the assumption that visitors would progress through it in a predetermined order), it also sheds some light on the so-called exhibition laws developed as a result of the early behaviourist studies undertaken by Melton and others (see Chapter 2).

Routes through the exhibition

The time data for each assembly visited by each subject were encoded in the order in which subjects visited them. It was possible, therefore, to map the different routes through the exhibition. Figure 14¹ shows the percentage of visitors making various successive moves through the exhibition, and these figures may be interpreted like probabilities. For example, the probability of taking the route A ('Cells') → B ('Growth') → BE ('Chromes') is .11 and A ('Cells') → BE ('Chromes') → B ('Growth') is .25. Although calculations of successive moves as presented in Figure 14 give the most accurate picture of particular pathways through the exhibition, the sample soon fragments into smaller groupings, indicating just how idiosyncratic individual movements are. (It is interesting to note that the exhibition was designed on the assumption that visitors would follow a sequence A → B →

¹. See page 90 for description of sections (assemblies) in the Hall of Human Biology.



* Less than 1%

Fig. 14. Routes through the Hall of Human Biology.

C → D → E → F → G → H → I) Another way of looking at routes and ignoring the turbulence represented in Figure 14 is to calculate the sequences of 'previous moves' and 'next moves' across the sample. That is, given that a subject has arrived at an assembly it is possible to calculate which assembly he was at previously; and conversely, given that a subject has arrived at any one assembly, it is also possible to calculate where he goes next. 'Previous moves' and 'next moves' for the sample of visitors is presented in the form of an origination/destination matrix in Table 16. The rows and columns in the matrix should be read separately. For example, to discover which destinations visitors were observed to go to from A ('Cells'), reference to the row in the matrix showing A, reveals that 240 visitors were observed leaving A. Of these, 124 went to B ('Growth'), 84 went to BE ('Chromes'), 2 were 'lost' while their behaviour was being recorded and a further 30 left the exhibition (shown as 'END' in the matrix). Reference can be made to the columns in the matrix to discover which assemblies visitors were in, immediately before arriving at the assembly referred to. For example, reference to the column E ('Homstis'), shows that 219 visitors were observed arriving at E. Of these, 130 came from D or DE ('Control'), 46 from F ('Hormes'), 36 from C ('Move') and 7 from B ('Growth'). These data are represented diagrammatically in Figure 15 which can be read in the same way as the data in Table 16.

'Next moves' can also be expressed as percentages on the row totals and when this is done, they may again be treated like probabilities of the various possible routes. These are presented as departure probabilities in Table 17.

The route A ('Cells') → BE ('Chromes') → B ('Growth') is $2\frac{1}{2}$ times as popular as A ('Cells') → B ('Growth') → BE ('Chromes')

Table 16.

Movements from assembly to assembly

DESTINATION

ORIGIN	DESTINATION															
	A ('Cells')	B ('Growth')	BE ('Chromes')	C ('Move')	D & DE ('Control')	E ('Homstis')	F ('Hormes')	FE ('Sexhormes')	G ('Hormnerv')	H & HE ('Mem')	I ('Percep')	J ('Cogdevel')	Bookshop at end of exhibition	Subject lost during recording	END	Totals
A ('Cells')		124	84											2	30	240
B ('Growth')	51		52	141		7								3		254
BE ('Chromes')	39	94		1											4	138
C ('Move')		31			139	36						1		1	6	214
D & DE ('Control')		1		37		130	1						1	1	3	173
E ('Homstis')		31	1	33	32		148			1					1	219
F ('Hormes')					1	46		40	18	73	27	48			3	257
FE ('Sexhormes')							6		3	76						85
G ('Hormnerv')							3	20		20	1	1				45
H & HE ('Mem')							49	24	24		69	10	0	2	6	184
I ('Percep')							14			11		72	2	1	19	119
J ('Cogdevel')				1			35			2	17		18		71	144
Bookshop at end of exhibition	3			1							5	10			2	21
Lost	1		1		1		1	1		1		2			2	10
Totals:	94	254	138	214	173	219	257	85	45	184	119	144	21	10	147	

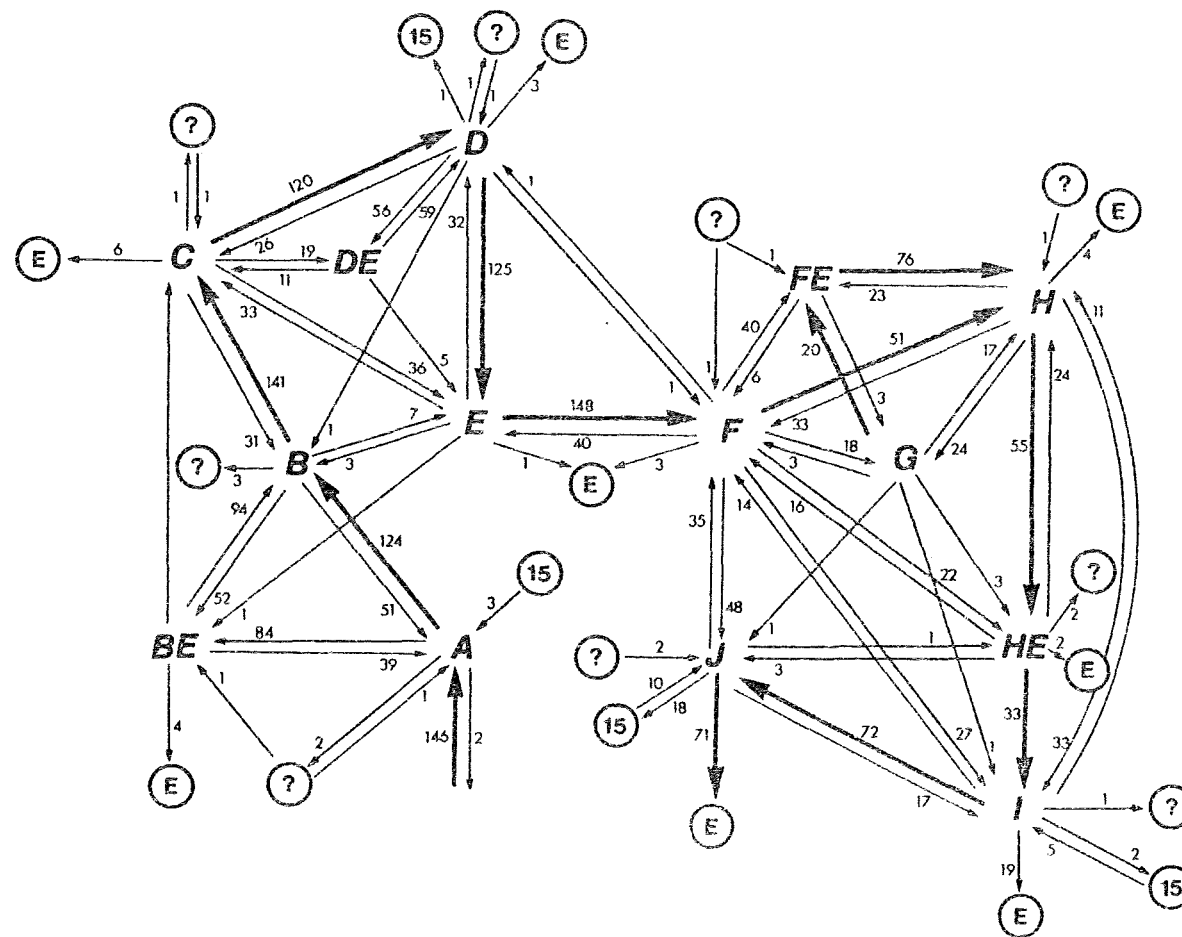


Fig. 15. Diagrammatic representation of movements from assembly to assembly (see Table 14)

Table 17

Departure Probabilities
(expressed as percentages)

A $\frac{13}{-}$ End		13	
A $\frac{52}{-}$ B $\frac{56}{-}$ C	29	50	
A $\frac{52}{-}$ B $\frac{20}{-}$ BE	10		
A $\frac{52}{-}$ B $\frac{20}{-}$ A	10		
A $\frac{35}{-}$ BE $\frac{28}{-}$ A	10	34	
A $\frac{35}{-}$ BE $\frac{68}{-}$ B	24		
		<u>97</u>	
B $\frac{3}{-}$ E	3		BE $\frac{3}{-}$ E 3
B $\frac{20}{-}$ A	20		BE $\frac{68}{-}$ B 68
B $\frac{56}{-}$ C	56	93	BE $\frac{28}{-}$ A 28
B $\frac{20}{-}$ BE $\frac{68}{-}$ B	14		<u>99</u>
B $\frac{20}{-}$ BE $\frac{1}{-}$ C	-		
B $\frac{20}{-}$ BE $\frac{28}{-}$ A	6	6	
		<u>99</u>	
F --- E	18	27	
F $\frac{16}{-}$ FE $\frac{7}{-}$ F	1.12		
F --- FE $\frac{4}{-}$ G	0.6		
F --- G	7		
F $\frac{20}{-}$ H	20	43	
F $\frac{16}{-}$ FE $\frac{89}{-}$ H	14		
F $\frac{9}{-}$ HE	9		
F --- I	11	30	
F --- J	19		
		<u>100</u>	

and almost as popular as A ('Cells') → B ('Growth') → C ('Hormes') though unlike the other two, it does not accord with the sequence of ideas in the exhibition. (These results are comparable to those obtained from the 'successive moves' data in Figure 14.)

The destination when leaving F ('Hormes') is as likely to be H ('Mem'), J ('Cogdevel') or E ('Homstis') as FE ('Sexhormes'), and two to three times as likely to be any particular one of these as to G ('Hormnerves'). The routes after F ('Hormes') are more varied than those before it. Only 27% of visitors follow the intended sequence of ideas - more people, 30%, move directly to I ('Percep') or J ('Cogdevel'). The most favoured destination is H or HE ('Mem') with 43% of the sample going here, direct or via FE ('Sexhormes'). This is not very surprising given the layout of the exhibition. When the visitor leaves F ('Hormes') he can, literally, move in any direction. Up to this point, fewer options were open to him as he moved between assemblies.

The routes taken by visitors through the Hall of Human Biology are of immediate interest to planners of the exhibition in that they show, generally, visitors do not follow the sequence in which the ideas are developed and presented according to 'learning hierarchy' notions (see Chapter 1). Moreover, as mentioned earlier, they have a bearing on the results of earlier studies.

For example, one of the supposedly invariant findings arising from all previous American studies, namely the pronounced right turning tendency, has not been found in the study. The first clear opportunity for visitors to take a right turn is as they leave Assembly B ('Growth'). At this point visitors can proceed straight ahead and progress to C ('Move'), as indeed the majority do, or they can take a left out of the exhibition or turn right into

E ('Homstis'). The percentage of visitors turning right into Assembly E was two per cent. The next clear opportunity to take a right turn (without back-tracking) is upon leaving Assembly F ('Hormes'). Here, visitors can turn right into J, left into FE or E, or they can proceed straight on into H. In fact, the destination when leaving F is as likely to be H, E or FE as J.

The present study also failed to find the so-called 'exit effect', that is where exits are supposed to be more attractive than exhibits in attracting visitors. It is possible for visitors to leave the exhibition from A, BE, C, D, E, F, H, HE, I and J. J ('Cogdevel') is the assumed exit, and most visitors do in fact leave by J. The remaining exits are clearly less popular than the exhibits. A sizeable minority do leave the exhibition from I, but there is an exit from I which if taken by visitors unfamiliar with the layout of the exhibition (which the overwhelming majority - being first-timers - are) gives him or her an impression of having come to the end of the exhibition rather than of having left it prematurely.

Furthermore, there is no evidence of museum fatigue, as conceptualised by Melton, in the Hall of Human Biology. Figure 15 shows that the modal path through the exhibition is $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E \rightarrow F \rightarrow I \rightarrow J$. The average times spent by visitors in I and J are longer than any other assemblies and the various measures of attracting power show that although progressively fewer visitors remain in the exhibition from A to J, those that do stay do not become attracted to fewer and fewer exhibits along the way.

It has not been possible from the present study to discover if prior interest had any noticeable effect on the routes taken through the exhibition. Unfortunately, at the time the study was undertaken,

visitors leaving one assembly for another had no means of knowing which assembly they were in until they were in it. Consequently, if a visitor discovered he was in an assembly in which he had little or no prior interest and as a result left it, it would not be possible to distinguish such a visit from a visitor who remained in the same assembly because he had a high prior interest in it.

Discussion

The results from the observations of visitor's behaviour have indicated which exhibits are attractive (that is, have high attracting powers) and which are not so attractive, but the reasons they are so have not been convincingly elucidated by the notion of prior interest. Although some evidence has been forthcoming which suggests that prior interest does influence whether or not an individual is attracted to an exhibit, it is also clear that other factors must be at work in attracting visitors to exhibits. What these factors are precisely still remains an open question and one which has been tackled by the two studies to be reported in the next chapter. Given the knowledge gained from the present study, all that can be said is that there must be other factors influencing the dynamics of the visiting process causing visitors to stop at certain exhibits and pass others by. It is this empirical question which has been assayed in studies reported in the next chapter, where an attempt was made to elucidate the perceived characteristics of exhibits and relate them to the attracting power of exhibits (as measured by the study reported here), thereby providing an explanation of what makes an exhibit attractive; and, conversely, what makes one unattractive.

However, on their own, the results from the present study have practical implications in so far as museum planners are concerned

to mount exhibitions dealing with topics of interest to their visitors. Clearly, the range of possible exhibitions a museum might mount is far greater than the finite number that can ever be accommodated in the space available. This study has shown that it is possible for visitors to express their interest in different topics and that the interest is related to an exhibitions' effectiveness as measured by the attractiveness of the exhibition, defined and measured behaviourally. Although weak, the relationship between prior interest and attractiveness was significant, which does suggest that exhibitions developed to match the interests of visitors would increase the possibility that visitors might learn from them. (This is particularly important if the philosophy of the planner is primarily didactic.) All previous studies have shown that casual visitors 'learn' nothing from exhibits (see Chapter 2) so that even a marginal increase in their potential effectiveness would be an improvement.

Finally, it is timely to re-emphasise the implications from these results to the design of educational exhibits organised along the lines of Human Biology. The Hall of Human Biology was designed around a learning hierarchy which assumed that visitors will follow a fixed sequence and look at all the exhibits for an amount of time necessary to receive their intended messages. There is abundant evidence from the study to show that visitors do not follow the intended sequence nor do they spend sufficient time at exhibits to comprehend the intended messages.

Despite these misgivings, visitor surveys at the Museum (see Appendix 1) have shown consistently that since its opening, the Hall of Human Biology is the most popular exhibition in the Museum among all types of visitors. It is clear that visitors can comprehend enough for their own purposes, even if their responses to exhibits are not what the designers of the exhibition had hoped for.

CHAPTER 7

CHARACTERISTICS OF IDEAL EXHIBITS

Introduction

The study reported in the previous chapter failed to give a satisfactory account of the differential attractiveness of exhibits in terms of prior interest. In the present chapter, two studies (named Study I and Study II) are described in which an attempt was made to provide a more complete account of the dynamics of the interaction between visitors and exhibits by examining the ways in which visitors perceive exhibits, and by relating these percepts to the exhibits' attracting powers as computed in the last chapter's study.

In Chapter 3, it was argued that museum exhibits could be conceptualised in terms of prototype ideal exhibits and that real exhibits varied in their distance from the ideal. Behaviourally, the further away an exhibit is from the ideal, the fewer people it will attract and the closer to the ideal, the more people it attracts. In abstract, the ideal exhibit can be thought of as that exhibit which attracts all visitors; and theoretically it is, therefore, the 'limit' to which real exhibits approximate in varying degrees.

Real-world exhibits will possess perceived characteristics with both positive and negative valences, and the higher their overall valence, the higher their attracting power. It follows, logically, that the more characteristics a real-world exhibit has in common with the ideal, the higher will be its attracting power since the attracting power of the ideal exhibit equals unity.

However, ideal exhibits do not exist as real-world objects nor do they necessarily exist as images in the minds of individuals. Accordingly, it might be difficult for an individual to describe what an ideal exhibit would look like. Nevertheless, it is a relatively straightforward task for an individual to state which characteristics the ideal should possess if he is presented with a list of characteristics to choose from. (For example, should the ideal exhibit be 'boring', 'exciting', 'easy to understand', etc.) Ideal exhibits can thus be conceptualised as the optimum combination of characteristics associated with real-world exhibits, yielding the highest possible valence and attracting all visitors. Consequently the concept of an ideal exhibit is as a purely theoretical construct against which real exhibits can be measured in terms of their proximity to it.

It is an empirical task to elicit the characteristics of real exhibits and then to measure the extent to which real exhibits and a putative ideal possess them. The final studies to be presented in this thesis comprised two separate, though related pieces of research concerned with this task. In the first of the studies, the aim was to elicit from visitors an exhaustive list of characteristics used by them to describe and discriminate between real exhibits. In the last study reported, the purpose was to measure the extent to which real exhibits and a putative ideal exhibit possessed each of the characteristics elicited in the first study reported in this chapter; and further to see the extent to which real exhibits approximated to the ideal.

There are over one hundred and twenty exhibits in the Hall of Human Biology and to reduce this number to manageable proportions for studying, a sub-group of forty five exhibits was chosen. Exhibits in the exhibition differ markedly in terms of the morphological characteristics, for example, models, slide-shows, interactive

games, text panels and so on. The exhibits selected for the studies were broadly representative of the different morphological types and the numbers from each section were roughly in proportion to the total number of exhibits in each section. In addition, the chosen exhibits were representative of the range of values of 'attracting power' and 'holding power' of all exhibits obtained from the unobtrusive study of the behaviour of a sample of visitors reported in the previous chapter (see Figures 12 and 13). The list of exhibits together with their respective attracting powers and holding powers is given in Table 18. The exhibits are listed in random order and the numbering is used throughout this chapter when referring to the exhibits. In addition, photographs of the majority of the exhibits can be seen in Appendix III.

Study I

A number of different methods are possible to elicit the perceived characteristics associated with the exhibits. It was decided to take subjects back into the exhibition to the exhibits after they had left the exhibition at the end of their visit. By this time, subjects would have formed an overall impression of the exhibition and made comparisons - consciously or not - between the various exhibits they had encountered. Subjects were taken to exhibits to elicit characteristics rather than asking them to recall exhibits or use visual prompts. The reasons for this approach have their origins in a consideration of the attitude concept presented by Harré and Secord (1972). They draw the distinction between 'avowed' attitudes (typically measured by pencil and paper tests in low arousal situations) and attitudes which are manifested in 'hot' situations, at the point of action, so to speak. It seems logical to expect different expressions of attitude or opinion towards the same object in these different situations. Taking subjects to the exhibits places them in a genuine context

Table 18

Exhibits selected from the Hall of Human Biology for studies of exhibit characteristics.

Exhibit	Attracting Power	Holding Power
1. B6 More About the Placenta	0.18	8
2. I9 Knowing How Large It Is	0.33	55
3. F1 Emergency	0.82	55
4. BE6 Boy or Girl - What Decides	0.19	18
5. H2 We Learn Skills	0.56	85
6. D7 The Brain Gathers Information	0.77	35
7. H5 We Store Our Experience as Memories	0.68	35
8. A6 The Sperm	0.52	18
9. H3 We Learn by Organising Our Experience	0.38	85
10. G1 Hormones and Nerves	0.35	13
11. FE1 Sex Hormones Before Birth	0.07	13
12. B9 Basis of Growth is More Cells	0.42	8
13. I22 Perception or Imagination	0.07	13
14. C5 Human Skeleton	0.10	8
15. J3 The Smell Game	0.43	65
16. J14 Alice's Tea Party	0.46	75
17. E4 Homeostasis Machine	0.40	45
18. A4 Living Cells	0.33	25
19. A1 The Living Cell - Cell Structure	0.61	18
20. I19 Perception of Others	0.42	65
21. D2 The Outer Layers of Your Brain	0.60	13
22. D5 Neurons: Networks and Layers	0.43	25
23. DE3 Sensory and Motor Figures	0.41	35
24. C3 Take Away the Bones	0.57	8
25. F10 Feedback	0.28	18
26. C7 Joints	0.57	18
27. J12 Mental Images	0.08	18
28. J23 John's Four Journeys	0.35	65
29. C2 Pull This Lever	0.53	13
30. J22 A Problem in Proportion	0.32	85
31. BE7 A, B, AB or O	0.20	25
32. A10 Fertilisation	0.47	18
33. C4 Where Would You Find These in Your Body	0.33	8
34. I23 Seeing Patterns	0.19	13
35. J25 Find the Combination	0.13	8
36. F6 Hormones and Glands	0.67	45
37. J17 The Tharks Have Landed	0.54	35
38. B3 Womb Area	0.75	35
39. B1 Life Before Birth	0.65	25
40. J15 Necklace Game	0.39	35
41. E5 Homeostasis	0.20	18
42. C14 Sending Signals	0.40	25
43. I4 Associating Sight with Sound	0.54	35
44. D9 Ideas and Decisions	0.32	8
45. I12 Recognising Things	0.58	35

of commitment quite unlike 'quiet' interviews outside the exhibition.¹

The aim of this study was to generate an exhaustive and salient list of attributes which could be communicated to all visitors without ambiguity. This latter point is particularly important since different meanings can be communicated by the same configuration of words. For example, Cherry (1961) has pointed out that the sentence "Do you think that one will do?" has a variety of meanings, depending on which word is stressed.

Method

Subjects

Twenty subjects above the age of 16 years, were selected systematically as they were leaving the Hall of Human Biology using the sampling method described in Chapter 6 and shown in detail in Appendix I.

Procedure

Each subject was taken back into the exhibition and invited to study three exhibits, in turn. After he or she had studied the first of three exhibits an elicitation interview (see below) was carried out. Upon completion of the interview, the subject was taken to a second exhibit and a further elicitation interview was carried out, whereupon completion the subject was taken to a third exhibit and the elicitation procedure was repeated again. The ordering of the

1. It will be remembered that prior interest as it related to the Chapter 6 study was measured in a low arousal situation before subjects entered the exhibition by asking them to rate their interest in seeing sections of the exhibition described to them in writing. Once in the 'hot' exhibition situation the avowed expressions of prior interest had only a weak effect in determining the attractiveness of the exhibition.

exhibits for subject's elicitation interviews was 'bias free' and was determined as follows. The forty five exhibits were arranged in random order. The first three were seen by the first subject, the second three by the second subject, and so on up to the fifteenth subject. A random sample of fifteen exhibits from the forty five was arranged in the same way for the last five subjects. The following elicitation was asked at each exhibit:

"Please tell me anything that strikes you particularly about the exhibit."

Verbatim answers were recorded and each subject was then prompted:

"Please tell me anything else that strikes you particularly about the exhibits."

Again verbatim answers were recorded and the prompt question repeated until the subject had exhausted his or her repertoire of comments about the exhibit. A laddering technique (Fransella and Banister, 1977) was used to elicit characteristics of a higher level of abstraction whenever subject's comments referred to the physical characteristics of an exhibit. For example, if a subject commented that an exhibit was red, the following question was asked:

"Yes, but what does the fact that the exhibit is red mean to you?"

When interviewing was completed, each individual verbatim answer and the exhibit to which it applied was written on a single card. The cards were then sorted into a number of piles, each pile containing cards covering conceptual categories judged to be similar by the experimenter. Next, a sub-list of items representative

of all the items in each of the categories was compiled. This list was then discussed with a further sample of twenty subjects to eliminate any semantic ambiguities.

Finally, the sub-list was reduced to a final list of approximately fifty items. This was achieved by giving the sub-list to three people who had worked on the design of the Hall of Human Biology with instructions to each of them to reduce the sub-list to a list of approximately fifty items by working independently. After they had each constructed their own lists they assembled together with instructions to reach a consensus about the final list of fifty or so items.

Results

Overall, approximately five hundred verbatim answers were recorded, written on cards and subsequently sorted into the following descriptive categories: Attractiveness/Noticeability, Overall Evaluation, Clarity and Ease of Comprehension, Evaluation of the Subject Matter, Required Visitor Response, Emotional Reactions, Visual Effect, Appeal to Different Age Groups. From these categories a list of approximately seventy items was compiled and this was further refined to produce the final list of forty eight items. The items are shown in Table 19.

Discussion

As already stated, the primary purpose of this study was to generate an exhaustive and salient list of characteristics to serve as an input to the final study reported next in this chapter. As such, it was essentially qualitative in nature; and it would be rash to draw any general conclusions about the differences between exhibits based on these findings alone. Nevertheless some

Table 19

List of Characteristics describing Exhibits in the Hall of Human Biology

Attractiveness/Noticeability

You can't help noticing it.
It's badly placed - you wouldn't notice it easily.
Your attention is distracted from it by other displays.

Overall Evaluation

If you arrive when one or two others are looking at it (or listening to it) you can't see it or hear it properly.
It's a memorable exhibit.
It makes you want to find out what it's all about.
It makes the subject come to life.
It makes learning easy.
It's part of a well laid-out section.
It arouses interest in the subject.
It's repetitive.
It's too long-winded.
It doesn't credit you with much intelligence.
It could make you look stupid in front of others.
It doesn't give enough information.
It presents information I'm already familiar with.
It assumes you know something about the subject already.
It teaches without being too serious.
It has taught me something new.

Clarity and Ease of Comprehension

Having seen it (or listened to it), it makes you think 'so what?'
The information is clearly presented.
It's confusing.
It deals with the subject better than textbooks do.
It gets the message across quickly.
It's difficult to know what point/s the exhibit is getting at.
It makes a difficult subject easier.
You can understand the point/s it is making very quickly.
It gets the message across clearly.

Evaluation of Subject Matter

It's a difficult subject to understand.
It deals with a complicated subject.
It deals with a subject of particular interest to me.
It uses familiar things or experiences to make the point.

Required Visitor Response

It takes a long time to see it (or listen to it) completely.
It allows you to test yourself to see if you are right.
You have to spend your time reading a lot of instructions telling you what to do before knowing what the exhibit it supposed to be about.
It's not clear what you're supposed to do or how you should begin.
You can take it at your own pace.

Emotional Reactions

It's entertaining.
It's artistic.
It involves you.
It's good fun.
I find it difficult to relate it to things or events in my life.

Visual Effect

It looks like something out of a fair-ground or fun-fair.
It looks like something out of science fiction.

Appeal to Different Age Groups

There's something in it for all ages.
It's more suitable for adults.
It's more suitable for young children (under 12).
It's more suitable for older children (over 12).

tentative conclusions have been drawn concerning the types of comments visitors made about the exhibits.

Firstly, there were very few comments made about situational factors. Comments were made about the positioning of exhibits (immediately noticeable or badly positioned so that they would not be noticed); about whether attention was distracted from an exhibit by other exhibits in the vicinity and whether or not an exhibit was easy to view when other people were already at it. Of course, this may have been an artefact of the elicitation procedure. Subjects were asked specifically about the exhibits themselves and not directly about exhibits in comparison with others or within the overall context of their visit to the exhibition.

A number of comments were elicited concerning the 'motivational power' of exhibits, that is, the power of the exhibits themselves to stimulate an interest in the topics being displayed. On the other hand, it seemed that certain exhibits had the opposite effect - for a variety of reasons they inhibited the visitors' interest in the topics displayed.

A great many descriptions were forthcoming about the quality and quantity of the information presented by exhibits. Many of these concerned the clarity of the exhibits' exposition and the amount of time it took to assimilate the information presented.

The relevance of prior interest was echoed in comments concerning subjects' personal interest in the topics displayed by exhibits; and other comments on the nature of the topics presented centred on the complexity of the information, per se, as opposed to the complexity of its presentation.

The interactive exhibits used for the elicitations generated a

number of comments specifically concerning the 'nature' of the interaction expected of the visitor rather than about the quality of the information or experience that was achieved as a result of the interaction.

Finally, there seemed to be a general concern among subjects that the exhibits should be able to make contact with a wide variety of visitors of all ages. These views were expressed by subjects without children as well as those with them.

Study II

The purpose of this study was threefold:

- (1) using the list of items generated in Study I, to measure the perceived appropriateness of the items to the forty five exhibits and a putative ideal exhibit;
- (2) to discover the similarities between exhibits and their distances from the ideal in terms of the relative attributes of the items; and
- (3) to investigate the relationship between the exhibits' proximity to the ideal and their attracting and holding powers.

The list of items generated in Study I is long and contains some redundancy by virtue of the fact that the descriptive categories (Table 19) contain different statements whose meanings appear to be somewhat similar. The experimental procedure in the present study would demand a shorter commitment of visitors' time if the list of items in Table 19 could be reduced without any loss of substantive information. Therefore, a fourth ancillary purpose was added to the study. This was to attempt an empirically systematic reduction of the list of items without fundamentally altering the relationships in the data obtained when the complete list is used.

Before describing the study in detail, a number of points need to be made. Typically, psychological studies which attempt to represent the similarities between a set of stimuli have involved techniques of multi-dimensional scaling (for a review of M.D.S. techniques see O'Hare, 1980) or cluster analysis (see Everitt, 1974) or a combination of both (for example, O'Hare, 1979). In

using MDS and clustering in psychological studies several different methods have been used to collect data but essentially they have all involved subjects making a judgement of similarity or dissimilarity between stimuli; and similarity and dissimilarity are assumed to be perfectly symmetrical. Tversky (1977) and Tversky and Gati (1978) have shown empirically that different results can be obtained when subjects are asked to focus on certain features of the stimulus objects rather than others. Furthermore, interpreting the results produced by MDS and clustering techniques is often speculative since both techniques represent relationships between stimuli in terms of overall similarity and not in terms of particular characteristics they are perceived to have in common.

In the present study, we were concerned to discover which exhibits resembled each other in terms of the relative attributions of the items (Table 19) to the exhibits by visitors to the exhibitions. Empirically, this involves asking subjects to attend to particular features of the exhibits (i.e. the items in Table 19) and make judgements about their presence or absence and not make judgements about overall similarity. This is much more in line with Tversky and Gati's contrast model concerning the way people perceive similarities and differences.

Before the study commenced, a small number of additional items relating to the physical characteristics of exhibits were added to the list given in Table 19 at the specific request of Museum staff who had worked on the development of the Hall of Human Biology. These were characteristics which staff felt had a particular bearing on the manner in which visitors behaved towards exhibits. They were:

It doesn't move, It's brightly lit and It's solidly built.

As will shortly be described, subjects were asked to state which of the items applied to the exhibits, scoring '1' if an item was thought to apply and '0' if thought not to apply. The use of such a heterogeneous list of items presents special difficulties from the point of view of reliability. Analysis of variance models of reliability (Maxwell and Pilliner, 1968) are not appropriate in this instance since they are concerned with items, all of which are supposed to measure the same quality. Theoretically, if we let the i th subject have the probability π_i of assigning an item to an exhibit, the data then consists of n observations, $x_1 \dots x_n$ (all taking values '1' or '0') and there are n parameters, $\pi_1 \dots \pi_n$. The reliability in the analysis of variance sense is the variance of the π s and there is no means of getting an estimate of this without considerable retesting. A test, re-test approach is not appropriate since it would not yield a good enough estimate of π_i . The major problem lies in the fact that the list, as it stands, contains no known element of duplication¹. As a rough and ready check on reliability, therefore, one item which duplicated an item in Table 19, namely, "It's good fun" was also added to the list. A final item - "It's above the average standard of exhibit in this exhibition" - was also included in the list and its significance to the issue of reliability will become clearer when the results are discussed. A randomised list of items used in the study is given in Table 20; the numerical codes of the attributes in the table correspond to those used in reporting the results.

1. It will be remembered that an ancillary purpose of the research was to reduce the list of items without altering the fundamental relationships within the data. As will be seen later in the chapter, this involved dropping items among groups of highly concentrated items. Thus, the items that were dropped were regarded as replicators of the items retained and the results obtained by dropping them gave an indication of the reliability of the instrument.

Method

Design

Asking subjects whether each of fifty three items applies to each of the forty five exhibits (plus their ideal exhibit) is far too many questions (=2,438) to ask a single subject in a single session. for this reason it was decided to use a balanced incomplete block design.

Annual visitor surveys at the Museum (Alt, 1980) have shown that visitors will happily participate in an interview lasting about fifteen minutes without much danger of drop-out. Pilot interviews showed that it took visitors less than fifteen minutes to judge exhibits across approximately fifty items. Therefore, it was decided to use blocks of two 'treatments' (i.e. judgements of exhibits across the fifty or so items) plus one other which was always a subject's rating of his or her ideal exhibit.

For design purposes the blocks can be thought of as containing two treatments since all subjects were asked to rate their ideal after they had rated two 'real' exhibits. A balanced block design for forty five treatments arranged in blocks of two treatments requires $45 \times 44/2$ paired comparisons for every pair of treatments to occur once within some block. Therefore, a balanced design requires 990 subjects and every treatment occurs 44 times. In this study two complete replications were sought to allow each of the 990 pairs of treatments to occur twice.

Subjects

1980 subjects, above the age of sixteen years were selected systematically as they were leaving the Hall of Human Biology using a method similar to that used in all the other studies reported in this thesis (see Appendix I).

Table 20

Randomised list of items used in Study II

1. You can't help noticing it.
2. It takes a long time to see it (or listen to it) completely.
3. It looks like something out of a fair-ground or fun-fair.
4. Having seen it (or listened to it), it makes you think 'so what?'
5. It allows you to test yourself to see if you are right.
6. It's more suitable for adults.
7. It's a difficult subject to understand.
8. It doesn't credit you with much intelligence.
9. You have to spend your time reading a lot of instructions telling you what to do before knowing what the exhibit is supposed to be about.
10. It could make you look stupid in front of others.
11. It looks like something out of science fiction.
12. The information is clearly presented.
13. It doesn't move.
14. It assumes you know something about the subject already.
15. It deals with a subject of particular interest to me.
16. It's entertaining.
17. It doesn't give enough information.
18. It's artistic.
19. It uses familiar things or experiences to make the point.
20. It presents information I'm already familiar with.
21. It's confusing.
22. There's something in it for all ages.
23. It's a memorable exhibit.
24. It involves you.
25. It deals with a complicated subject.
26. It's good fun.
27. It's brightly lit.
28. It's more suitable for young children (under 12).
29. It makes you want to find out what it's all about.
30. It's not clear what you're supposed to do or how you should begin.
31. It deals with the subject better than textbooks do.
32. It has taught me something new.
33. It's more suitable for older children (12+).
34. It makes the subject come to life.
35. It makes learning easy.
36. It's badly placed - you wouldn't notice it easily.
37. It gets the message across quickly.
38. If you arrive when one or two others are looking at it (or listening to it) you can't see (or hear) it properly.
39. It's difficult to know what point/s the exhibit is getting at.
40. I find it difficult to relate it to things or events in my life.
41. It's solidly built.
42. It teaches without being too serious.
43. Your attention is distracted from it by other displays.
44. It's part of a well laid-out section.
45. It makes a difficult subject easier.
46. You can take it in at your own pace.
47. It's too long-winded.
48. It arouses interest in the subject.
49. You can understand the point/s it is making very quickly.
50. It gets the message across clearly.
51. It's repetitive.
52. It's good fun.
53. It's above the average standard of exhibit in this exhibition.

Procedure

All subjects were asked to judge two of the forty five exhibits across the items in Table 20 followed by their ideal. The 1980 blocks or pairs of treatments were allocated at random to the 1980 subjects and on half the occasions one member of the pair of 990 treatments occurred first and on the other 990 occasions it occurred second.

Subjects were taken to the first exhibit of their pair and invited to look at it for a minute. They were then handed a set of cards on which were printed the items in Table 20 and each card was preceded by the following question:

"In your opinion which, if any, of these descriptions apply to this display?"

The experimenter simply ticked the items mentioned by each subject on a code sheet fixed in a clipboard he was carrying. Two prompt questions were used. These were: "Any others?" and "Is that it, then?" There were ten cards in all identified by the letters A - J. Card A contained items 1 - 6, card B items 7 - 11, card C items 12 - 17, card D items 18 - 22, card E items 23 - 27, card F items 28 - 32, card G items 33 - 37, card H items 38 - 42, card I items 43 - 47 and card J items 48 - 53. After the subjects had completed the task, the experimenter took them to the second exhibit in their predetermined pair and the procedure was repeated. Subjects were then taken out of the exhibition whereupon they were asked the following question:

"Now I'd like you to imagine what you think is an ideal display for the Human Biology exhibition and tell me which, if any, of the descriptions on this card would apply to the ideal exhibit?"

Subjects were handed a card at a time and the question together with the two prompts occurred with each card. Before each exhibit was rated, the cards were shuffled before they were presented to subjects.

All data were entered directly from the code sheets to an Apple micro-computer. Data entry was verified by a 'double entry' procedure and any inconsistencies were referred to the original code sheets for clarification and, if necessary, correction.

Results

The first part of the analysis consisted of computing a 'percentage matrix' showing the percentage of subjects endorsing each item for each exhibit and this was carried out on the apple computer. Part of the complete data matrix is given in Table 21. The exhibit profiles shown in Table 21 are of five exhibits which were close to the ideal in terms of the relative attributions of the items (exhibits 3, 26, 38, 6 and 29) and five exhibits with dissimilar profiles to the ideal exhibit (exhibits 27, 22, 31, 4 and 13). (The complete data matrix is given in Appendix IV).

The percentage matrix then formed the input to a Biplot analysis which is a sort of principal components analysis of rows and columns of the matrix simultaneously (Gabriel, 1971). The biplot thus gives a simultaneous 'map' or ordination of exhibits and items allowing one to see at a glance which exhibits resemble each other (in the sense of which items apply to them) and which exhibits are dissimilar. It also shows which items apply strongly to which exhibits. Everitt (1978) has given an example of a biplot applied to psychiatrists' ratings of patients. In Everitt's example, the analysis showed which patients were similar to one another in

terms of the ratings; and which psychiatric symptoms characterised the various patients. In the present study, the components representing overall level of response for exhibit and item have been removed since the concern was to discover which items discriminated between exhibits; and, in particular, which items characterised the 'ideal' exhibit. Thus the biplot shows not, say, that item j applies particularly strongly to exhibit i , but that relative to the general level of all items on exhibit i , item j is particularly high. In the present case, the analysis is identical to a principal co-ordinates analysis of exhibits (see Everitt, 1978) with the first axis removed.

The output from the biplot analysis is shown in Figure 16. The quadrant location is of no significance and the circled numerals which represent exhibits can be read more or less as one would read a map; the absolute distance between them represents the similarity between exhibits in terms of their profiles across the fifty three items. Examination of Figure 16 shows that exhibits close to the ideal exhibit (numeral 46, marked with an asterisk for ease of reference) are exhibits 3, 26, 38, 6 and 9; and exhibits most distant from the ideal are 29, 27, 22, 31, 4 and 13.

The 'map' in Figure 16 also allows one to estimate how strongly an item (items are represented by non-circled numerals) applies to an exhibit. This is achieved by constructing straight lines (= vectors) from the origin to the exhibit and items in question. Then, the cosine of the angle formed multiplied by the product of the two vector lengths gives the 'relevance index' of the item to the exhibit. Generally, the smaller the angle between the exhibit and item vectors and the longer the vectors, the more strongly the item applies to the exhibit. By the same token, item vectors orthogonal to exhibit vectors imply that those items are unrelated to the

Table 21

The percentage endorsement of items for ideal exhibit and five exhibits 'close' to the ideal and five distant to the ideal.

ITEMS	EXHIBITS										
	Ideal	3	26	38	6	29	27	22	31	4	13
1	62	97	72	90	73	60	0	26	23	17	17
2	9	20	1	2	3	0	10	7	30	28	7
3	15	26	32	8	67	48	0	10	34	31	7
4	0	1	10	2	3	12	27	23	8	9	40
5	71	0	20	3	34	42	1	13	28	32	18
6	8	6	0	3	0	2	16	31	47	46	2
7	13	2	1	2	1	0	8	56	67	59	11
8	0	6	10	5	14	9	10	1	0	3	33
9	3	0	0	0	2	1	33	16	45	37	15
10	0	2	2	0	0	8	1	2	3	9	11
11	21	3	1	56	2	0	0	26	1	1	16
12	94	51	67	72	88	76	63	31	48	55	28
13	1	2	7	43	3	0	42	30	5	7	52
14	31	31	36	22	10	20	20	62	49	46	16
15	35	7	9	26	8	7	2	12	7	14	4
16	59	87	43	33	65	56	2	3	14	6	10
17	0	9	31	11	12	21	34	58	23	30	61
18	40	15	1	51	17	3	3	23	8	5	16
19	66	78	39	15	79	44	2	5	27	9	33
20	2	33	45	48	44	33	31	12	21	28	26
21	0	0	0	0	1	0	6	45	36	34	21
22	84	77	68	86	53	76	3	10	16	8	27
23	43	71	11	90	14	16	3	8	8	6	4
24	79	31	68	50	58	88	7	21	44	30	22
25	49	8	16	27	21	9	26	51	72	67	12
26	52	56	38	23	72	65	1	6	16	6	7
27	51	37	6	41	63	31	2	36	24	17	9
28	1	1	16	6	30	20	0	2	0	0	12
29	87	62	36	35	29	30	12	30	53	57	19
30	0	1	1	2	5	1	14	53	36	32	57
31	55	36	63	77	49	51	3	19	35	28	7
32	60	3	11	19	1	2	3	15	17	11	6
33	2	9	14	6	13	7	29	29	33	26	7
34	86	88	60	85	74	71	1	17	20	20	10
35	50	28	47	56	58	45	2	12	28	26	7
36	0	0	1	1	1	3	79	16	17	9	72
37	40	49	38	35	60	63	8	9	15	13	10
38	0	20	5	2	12	17	44	10	40	36	17
39	0	0	7	1	0	2	35	36	15	8	51
40	0	0	9	3	0	2	10	26	19	11	9
41	36	15	73	50	24	70	14	16	41	41	7
42	61	41	39	25	85	49	2	20	28	18	65
43	0	5	9	1	3	4	67	44	26	18	65
44	71	62	69	73	63	67	37	56	43	46	30
45	65	13	23	27	21	19	2	20	37	39	3
46	75	6	68	57	50	56	41	34	50	49	30
47	0	1	0	0	1	0	21	0	15	5	2
48	78	60	38	63	42	45	8	26	36	38	18
49	43	50	47	39	65	60	13	14	6	13	12
50	95	73	59	66	66	58	16	17	36	38	10
51	1	12	0	11	3	2	1	1	1	1	1
52	45	29	22	16	58	42	1	8	8	6	4
53	99	93	70	92	90	70	2	22	48	55	15

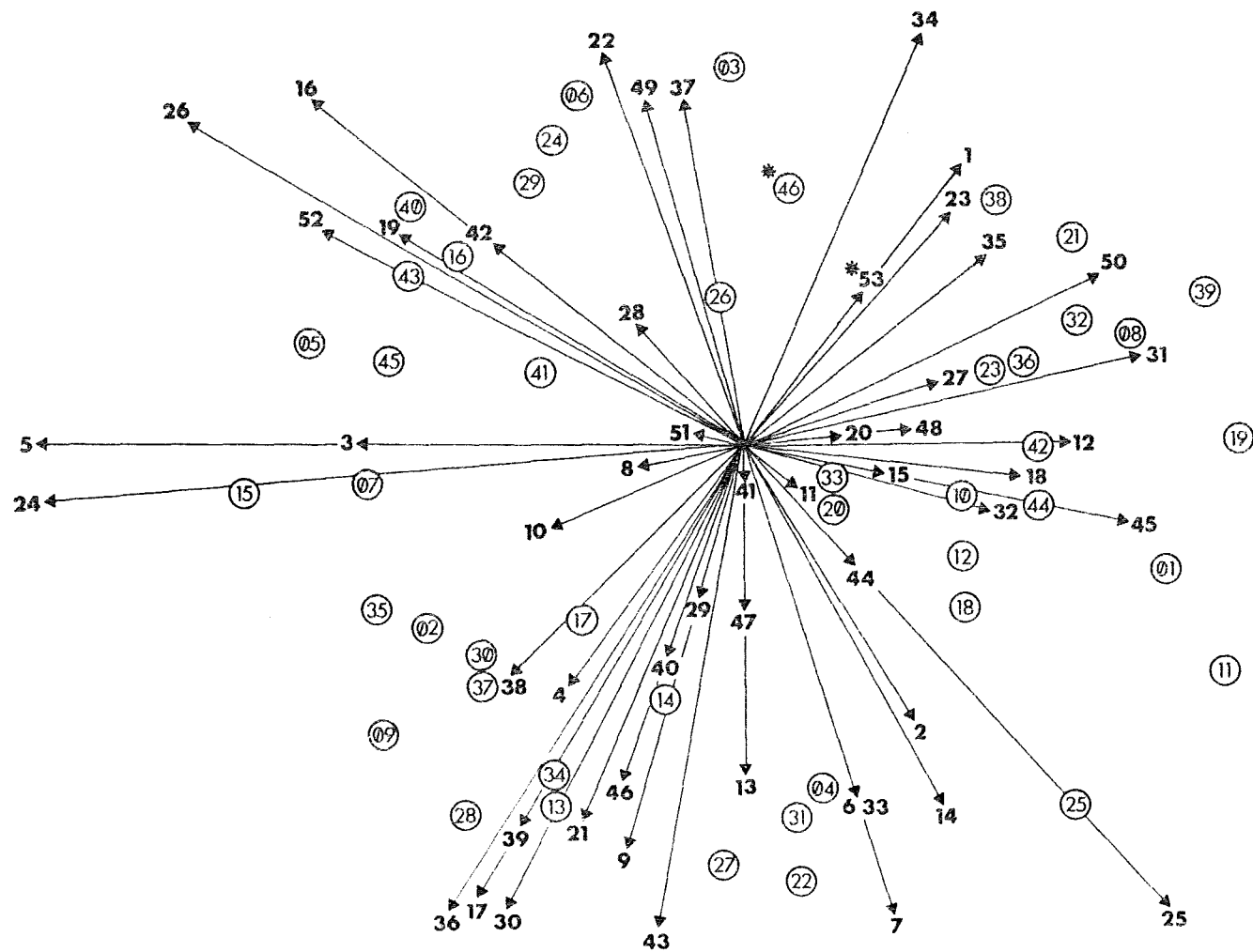


Fig. 16. Biplot of 46 exhibits and 53 attributes in the Hall of Human Biology (see text for explanation).

exhibits. Further as the angle approaches 180°, the relationship between the item vectors and the exhibit vectors becomes strongly negative. In the present case, this can be taken to imply that the opposite of the items 180° from the exhibit apply to the exhibit.

The items that described the ideal most strongly were 34, 37, 39, 22 and 1; and the items that are most strongly negative are 36, 17, 30, 43 and 7. For ease of reference these results together with items orthogonal to the ideal are summarised in Table 22.

Table 22

Relationship of items of Ideal Exhibit

Strongly Apply	Orthogonal	Strongly Negative
It makes the subject come to life	It allows you to test yourself to see if you are right.	It's badly placed - you wouldn't notice it easily.
It gets the message across quickly	It involves you.	It doesn't give enough information.
You can understand the point/s it is making very quickly.	It deals with a subject better than textbooks do.	It's not clear what you're supposed to do or how you should begin.
There's something in it for all ages.	The information is clearly presented.	Your attention is distracted from it by other displays.
You can't help noticing it.	It makes a difficult subject easier.	It's a difficult subject to understand.

Interpretation of the items in the strongly negative and strongly apply columns in Table 22 is straightforward but at first glance the orthogonal items seem to fly in the face of intuition and most certainly contradict the conventional wisdom currently extant among museum educators who have consistently asserted that participatory or interactive exhibits are exhibits to be aimed for in 'ideal' educational exhibits (for example, De Waard, 1974; Eason and Linn, 1975; Miles and Alt, 1979; Screven, 1968, 1970, 1974; Screven and Lakota, 1970 and Shettel, 1973.) The present research clearly suggests that there are a number of other requirements associated with the ideal exhibit of far greater importance than interaction or participation. Indeed, it may be that some of the characteristics of ideal exhibits as measured here are incompatible with the design of successful interactive exhibits. Certainly exhibits most distant from the ideal include among their number the most interactive exhibits to be found in the Hall of Human Biology.

Item 46 - You can take it at your own pace - does, however, seem anomalously placed so distant from the ideal but this can be explained by the structure in the data set. Exhibits at which items 36, 17, 30, 21, 9 and 43 were thought to apply strongly were also exhibits that visitors could take at their own pace. Therefore, by association, the 'best fit' for item 46 was to place it in its present position; in fact 75% of subjects thought that item 46 also applied to the ideal.

The duplicated items - 26 and 52 - appeared in almost the same direction but laterally displaced slightly indicating some measurement error. (If there was no error, 52 and 26 would be coincidental.) Item 53 appears in the same general direction as the ideal exhibit, lending empirical support to the concept of an

ideal exhibit for a population of visitors. The length of the exhibit vector, 46, also supports this view since if individuals were widely divergent in their perceptions of an ideal exhibit, it would not be strongly related to any item and so would appear near the centre of the plot. The length of vector 53 suggests that the item did not discriminate particularly well between exhibits and indicates the possibility of a response bias whereby subjects were somewhat reluctant to be critical of the exhibits they were asked to judge. As pointed out in Chapter 6, this sort of response bias has been found in almost all studies undertaken at the Natural History Museum when subjects have been asked to make evaluative judgements about exhibits (Alt, In Press); and inspection of the original data matrix shows that approximately three quarters of the exhibits were judged to have been above average.

Next, the distance d_x of each exhibit from the ideal was calculated, measuring the straight line between them on the two-dimensional representation in Figure 16; and the correlation coefficients between d_x and attracting power and d_x and holding power were computed. Pearson's r equalled -0.63 for attracting power, highly significant at $p = 0.01$, but not significant for holding power ($r = -0.09$). The relatively high value of r for attracting power is extremely interesting psychologically. It suggests that visitors carried out a preliminary appraisal of exhibits to decide whether or not to stop. This appraisal was done very rapidly and very critically in that visitors apparently were able to make an immediate assessment of the accessibility of the information in the displays.

Finally, an attempt was made to reduce the list empirically. Two criteria were set against which the success of any reduction was to be judged. Firstly, exhibits should maintain their positions relative to the ideal when factorised using the reduced list;

and secondly, the high correlation obtained between attracting power and distance from the ideal when computed on the full list should be upheld when computed using a reduced list. A twofold procedure was adopted for dropping items from the list. Initially, items with short vectors in Figure 16, that is, the central items on the plot, were omitted on the grounds that they did not discriminate between exhibits. Then, items in highly correlated groups of items were dropped. Judgement was exercised as to which of the items in the groups were to be retained on the basis of the degree of semantic redundancy contained within the group. The list was reduced to twenty three items beyond which any further reduction would have resulted in the loss of whole categories of descriptions. The distances of each exhibit from the ideal were computed across the reduced list of items and correlated with their distances computed on the full list using Kendall's τ . Kendall's $\tau = 0.85$, highly significant at $p = 0.01$. The configuration of exhibits obtained by the biplot using the full and reduced list of items are shown in Figures 17 and 18, respectively. Overall, it can be seen that the broad patterns were maintained using the reduced list. Furthermore $r = -0.70$ ($p < 0.01$) when computed for attracting power and distance from the ideal using the shortened list. The reduced list is shown in Table 23.

Two observations may be made concerning the results of the procedure adopted for reducing the original list of attributes. Firstly, the removal of attributes among highly correlated groups of attributes without altering the overall pattern of the results suggests that it is reasonable to assume that certain attributes acted as replicates of other attributes; and as such the original measuring instrument had a high degree of reliability (see Footnote on page 145). Secondly, since the reduced list of items yielded almost the same results without losing substantive information

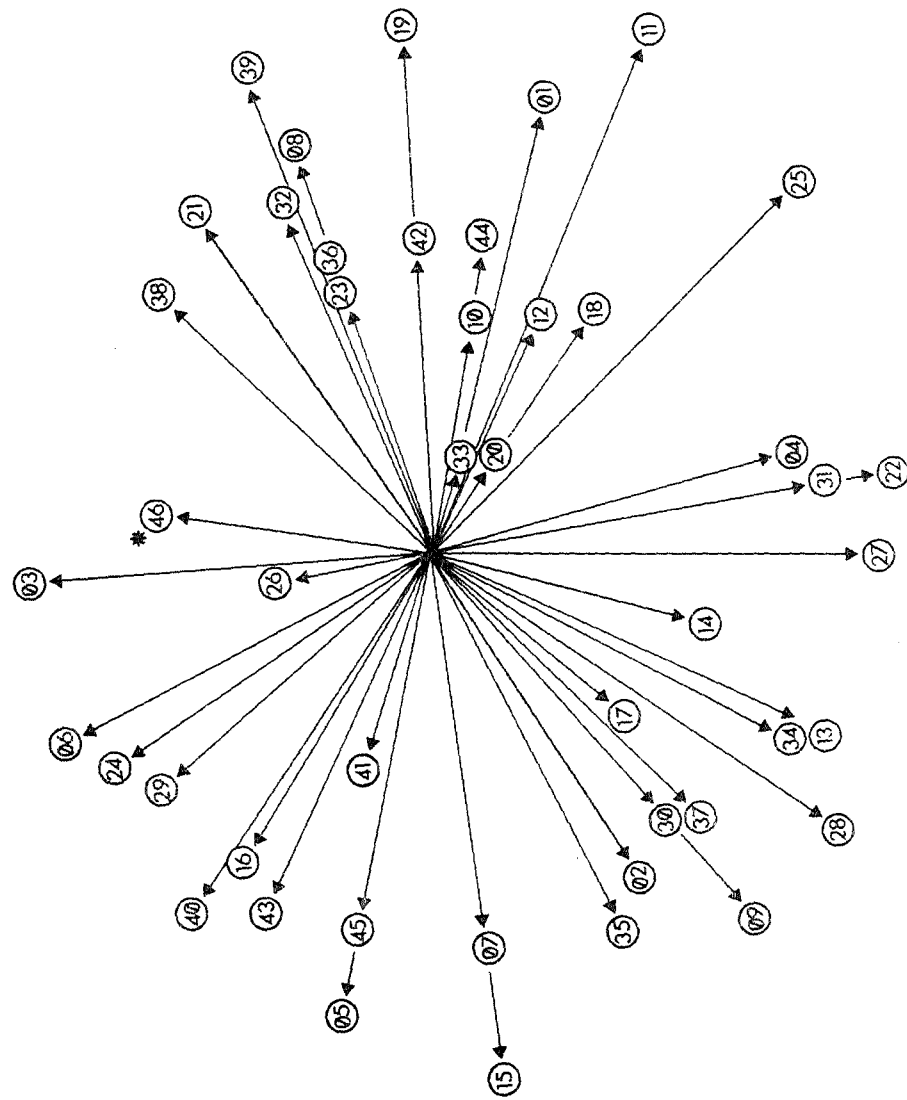


Fig. 17. Factorisation of exhibits across 53 items

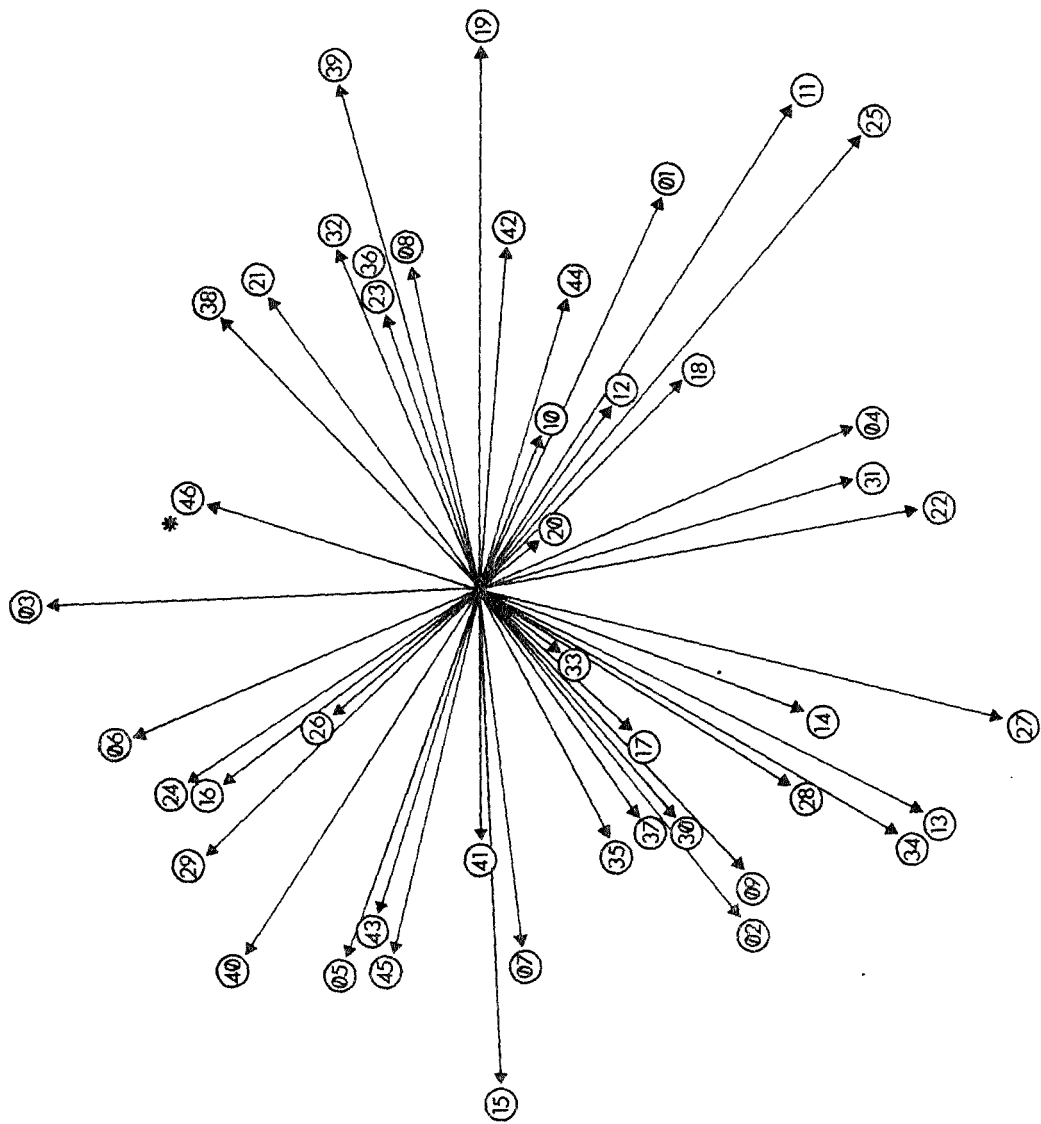


Fig. 18. Factorisation of exhibits across 23 items.

Table 23.

Reduced list of items from the original list of 53 items used for measuring the perceived characteristics of exhibits in the Hall of Human Biology.

It takes a long time to see it (or listen to it) completely.

It's a difficult subject to understand.

It doesn't move.

It doesn't give enough information.

It's artistic.

It uses familiar things or experiences to make the point.

It's confusing.

There's something in it for all ages.

It's a memorable exhibit.

It involves you.

It's good fun.

It's brightly lit.

It deals with the subject better than textbooks do.

It makes the subject come to life.

It's badly placed - you wouldn't notice it easily.

If you arrive when one or two others are looking at it (or listening to it, you can't see or hear it properly.

It teaches without being too serious.

Your attention is distracted from it by other displays.

It makes a difficult subject easier.

You can take it at your own pace.

You can understand the point/s it is making very quickly.

It's above the average standard of exhibit in this exhibition.

concerning the characteristics of exhibits, in any future study, it should be possible to use the reduced list, thereby reducing the amount of time required from subjects participating in the study. This is an important practical consideration since, after all, taking visitors away from exhibits which presumably they came to see to take part in surveys in an erosion of their limited time in the Museum.

Discussion

The results have provided evidence consistent with an approach in which real exhibits are conceptualised in terms of their distance from an ideal exhibit. An attractive exhibit is one that communicates a short clear message, which is displayed in a vivid manner and appeals to all age groups. More specifically, since a highly significant correlation was obtained between the attracting power of an exhibit and its distance from the ideal, the data have provided an "explanatory theory" of what makes an exhibit attractive.

These findings also call into question the validity of holding power as an index of exhibit effectiveness. Characteristics particularly associated with the ideal exhibit were those concerned with the rapid and clear transmission of a short message. Therefore, instead of attempting to capture visitors' attention in exhibits which convey a large amount of information (that is, attempting to generate high holding powers), designers should concentrate on attracting visitors with exhibits that deliver a short message. Furthermore, the absence of any correlation between distance from the ideal and holding power suggests that holding power is a construct which is unrelated to the manner in which exhibits are perceived and evaluated by visitors. It is, therefore, an index

which takes no account of the psychology of the Museum visitor.

Current theories about effective museum exhibits draw heavily on educational theory developed in mainstream education in which an influential dictum holds that the learner should be actively involved in the act of discovery, (for example Bruner, 1960; Dewey 1938 and Gagné and Briggs, 1974). While this may well be appropriate to the classroom it is not certain from these results that it applies to museums. It might, therefore, be a misguided leap of the imagination to equate the casual museum visitor with the classroom learner, and to think of the museum as a sort of informed school. It may well be wrong to think of the museum exhibit as an educational medium in the same sense as educational television, textbooks and programmed instructional material. Indeed, at first sight, the results from the present study appear to preclude the possibility of effective interactive learning through participatory exhibits; certainly by those that demand an unacceptably long call on the visitor's attention. However, since this conclusion seems to contradict the accepted notions of 'ideal' exhibits and are in some ways counter-intuitive, it would be rash not to consider them carefully.

Specifically, there needs to be a clear recognition that if an exhibit possesses attributes felt by visitors to apply strongly to the ideal exhibit, this does not imply it should not possess attributes found to be orthogonal or unrelated to the ideal. For example, it might be the case that if an exhibit could "get a message across quickly" and "it involved you", then the conjunction of these two attributes would result in a more desirable exhibit than one that simply gets across a quick message without involvement. It so happens that the overwhelming majority of exhibits in the Hall of Human Biology demanding involvement from the visitor are

also those that require him to invest a large proportion of his time in order to access the information in the exhibits. Because these participatory exhibits do not share many of the characteristics associated with the ideal exhibit, relative to the other items, 'involvement' does not emerge as a characteristic that applies strongly to the ideal. Nevertheless, inspection of the original data matrix (Table 21) reveals that 79% of subjects thought the item "It involves you" applied to the ideal exhibit. Thus, involvement is not an undesirable or even irrelevant characteristic. However what the results have established is a hierarchy of desirable characteristics for museum exhibits; and it is manifest that unless an exhibit promises to deliver a short, clear message it is unlikely to attract the attention of visitors. This is a salutary reminder to museum planners who implicitly appear to assume that visitors have an infinite amount of time at their disposal.

The methodology of the present study also provides a new approach to classifying museum exhibits in terms of their effectiveness by placing exhibits according to their proximity to an ideal exhibit. Previous attempts to classify museum exhibits and relate the classification systems to behavioural indices of exhibit effectiveness have failed because classification systems have been conceptualised in terms of the designers' purpose in designing exhibits (Shettel, 1968) or in terms of the different media used to represent information (Bretz, 1971), thereby failing to recognise the psychology of the museum visitor as an active agent who imposes his own cognitive constructions on the exhibits he perceives. Before turning to the final chapter in which an attempt will be made to connect the major arguments developed in the thesis and relate the ideas to more general issues, a few concluding comments will be made on the rationale underlying the experimental procedures

adopted, particularly as it related to the second study.

In reading the present chapter, there would be some justification if the reader came to the conclusion that there was an implied assumption underlying the studies to the effect that any one visitor could be regarded as the same as any other, regardless of age, educational background or any number of other possible subject classifying variables. It is acknowledged as self-evident that within such a heterogeneous population as casual visitors to an institution like the Natural History Museum, there will be a great many individual differences in interests, beliefs, expectations and so on. Furthermore, it would be absurd to deny that these differences have an influence on how the individual acts in the Museum. A more complete psychological description of the interaction between visitors and exhibits than the one given here requires detailed investigations of individual visitors to the Hall of Human Biology. A possible approach to this issue will be discussed in the next chapter.

Nevertheless, the theory of exhibit effectiveness outlined in Chapter 3 did take the individual visitor as its starting point and, in particular, his perceptions of the exhibits he encounters during his visit to the Museum. For mainly practical reasons, (after all, an evaluator is paid to provide practical suggestions) the emphasis was changed to consider a population of visitors. This was in the belief that if the theory was to be of any value to planners and designers of exhibits, it must at least be capable of accounting for a significant proportion of the variation within a population of visitors in order to bring the varied behaviour of individual visitors into a tractable focus, thereby providing a basis for practical action.

The theory was primarily concerned with explaining why certain exhibits succeeded in attracting visitors whereas others did not. Therefore, it was a fundamental prerequisite that a population of exhibits was studied to provide a basis for understanding the differential attractiveness within the population of exhibits. Accordingly, forty five exhibits representative of the distribution of attracting powers of all exhibits in the Hall of Human Biology were chosen for study. In the second study it was necessary to recruit approximately two thousand subjects to yield a total of eighty eight replications of the evaluations of each of the forty five exhibits. To have investigated the perceptions of the exhibits by different sub-samples of visitors (defined by various subject classifying variables) would have required far more than two thousand subjects to have yielded results at a reasonable level of precision. It was simply not practicable to have pursued this course at the time the studies were conducted.

In conclusion, the findings of the studies reported in this chapter are probably best regarded as a spring-board for more detailed investigations of individual visitors within the framework of the theory of exhibit effectiveness outlined in Chapter 3 of the thesis.

CHAPTER 8.

OVERVIEW

Summary and Conclusions

In chapter 1, this thesis began with a description of the Hall of Human Biology which was the first of several exhibitions to be opened as a consequence of a new exhibition scheme started at the British Museum (Natural History) in the early 1970's. As pointed out by Miles and Alt (1979) the exhibition scheme is new in three distinct ways:

- "1. in the scope of its content - it is thematic and highly innovative in concept;
2. in the way this content is to be presented to the public - the exhibits aim at being self-explanatory and at involving visitors actively in a number of ways;
3. in the way that the task of setting up the exhibits is being approached."

Of particular interest was the second of these three aspects since an attempt has been made to bring together and apply knowledge from a number of different areas, not least among them being the systematic evaluation of the effectiveness of the exhibits once developed and opened to the public. This outlook has been described by Miles and Tout (1979) as the scientific approach to the advancement of knowledge about effective exhibits; and they

characterised it formally in terms of the criticism-based schema presented by Popper (1972) as a three stage iterative process of trial and error:

$$P_1 \longrightarrow TT \longrightarrow EE \longrightarrow P_2$$

where P = problem, TT = tentative theory, EE = error elimination. Here, exhibits are regarded as trial solutions (tentative theories) to suitably defined problems and the four entities linked in these three ways were identified in the context of exhibition design.

The idea of evaluating museum exhibits to eliminate errors, thereby providing a knowledge-based approach to exhibition design is not new. There is now a considerable body of literature on the evaluation of museum exhibits generated mainly in America as a result of studies in museums and art galleries. These studies were reviewed in Chapter 2 of the thesis. Criticism was raised on three broad fronts:

1. methodological;
2. philosophical;
3. the utility value, or usefulness, of the research to designers of exhibitions.

In Chapter 2 it was shown that the enormous amount of research already undertaken has yielded little by way of useful practical guidelines to exhibit designers to help them in the design of exhibits. Although most of the studies reviewed in Chapter 2 suffered from considerable methodological limitations, particularly from the view of the validity and reliability of the measures of

effectiveness taken, it was suggested that these limitations were not the major reasons why the studies had failed to furnish designers with useful information. Fundamental criticism was aimed at the philosophical level. Sepcifically, it was pointed out that all so-called theories of exhibit effectiveness (implied or expressed) and the studies that had emanated from them are underpinned by a model of the museum visitor as an essentially passive creature. Accordingly, the behaviour of the museum visitor can be manipulated by exhibits without regard to his own interests and preferences. Thus, it has been assumed that visitors ought to behave predictably in given conditions and they can and should be manipulated by engineering appropriate conditions (i.e. effective exhibits). Learning at exhibits, the routes taken through an exhibition, exhibit interest, etc. have been conceived as dependent variables, that is as functions of exhibit design quite independent of the psychology of the museum visitor.

Over the years two indices have become established against which the effectiveness of exhibits are judged. These are:

Attracting Power - defined as the probability of stopping at an exhibit, that is, the ratio of 'stoppers' to non-stoppers, and

Holding Power - defined as the average (mean) time spent at an exhibit by those visitors who stop.

An effective exhibit is deemed to be one that attracts a large number of visitors who, moreover, remain long enough to receive the designer's intended message. While it is self-evident that an exhibit which fails to attract visitors cannot be effective, in Chapter 2 it was argued that the validity of 'holding power' as an

index of effectiveness was suspect if an active-dynamic conception of the museum visitor is upheld. All the studies reviewed have shown that the average holding power of exhibits is in the region of thirty seconds so that if the museum visitor is viewed as an active agent who is not prepared to invest long periods of time viewing individual exhibits, an effective exhibit would be regarded as one that transmits a short message. Earlier workers who have taken an essentially deterministic view of the museum visitor have maintained that an effective exhibit is one that manipulates the visitor to remain for the required time, apparently without an upper limit.

In Chapter 3 a new theory of exhibit effectiveness was put forward. Essentially, it takes as its starting point the idea that exhibits are in competition with one another for the attention of visitors. An effective exhibit is one that competes successfully and attracts visitors and ineffective exhibits are those that compete unsuccessfully. An explanation of what makes exhibits attractive is to be found in terms of their 'characteristics' as perceived by visitors. In other words it is, at least in form, a person-exhibit interactive model. Within such a framework certain characteristics will be evaluated favourably and others unfavourably. Characteristics that are perceived favourably have a positive valence (to borrow Lewin's terminology), while those that are evaluated unfavourably have a negative valence and will tend to repel visitors. For a given visitor the sum total of the different valences will determine whether an exhibit attracts or fails to attract that visitor. The concept of a 'putative' ideal exhibit was introduced as a theoretical construct and was defined as a notional exhibit that attracts all visitors, that is, has an attracting power of unity. Real world exhibits were construed, accordingly, in terms of their distance from the ideal. Thus the more characteristics a real world

exhibit shares with the ideal, the higher its attracting power. The remainder of the thesis was concerned with testing the theory empirically in the Hall of Human Biology.

The theory requires that the attracting powers of real world exhibits are measured and Chapter 4 describes the development of an apparatus for doing so. Initially, a method whereby an observer physically followed visitors as they wandered through the Hall of Human Biology, noting the exhibits they stopped at (and for how long), was attempted. This was the method that had been adopted by the majority of other researchers in the field. Unfortunately, when ever such attempts were made, the observer became conspicuous to the visitor being tracked. To overcome this difficulty, closed-circuit television was installed throughout the Hall of Human Biology to allow an observer to watch and record the behaviour of individual visitors. In addition, a real-time event-recorder based round a 'PET' micro-computer was developed. This permits the observer to record and analyse the behaviour of a visitor at the actual time he is observing, thus obviating the need for retrospective analyses of video recordings. This system is described in detail in Chapter 4.

Chapter 5 is devoted to a general discussion of the notion of reliability in observational research and in particular reported the findings of an inter-observer reliability study using the tracking-recording apparatus developed and described in Chapter 4.

Chapter 6 reported a major observational study of a sample of visitors to the Hall of Human Biology. Using a model developed by Miles and Tout (1979), the attracting power and holding power of exhibits were defined formally in terms of the interaction between visitors and exhibits in the form of a matrix. The model was extended to develop statistics pertinent to the behaviour of visitors

in addition to those defined for exhibits. Specifically, the 'attractiveness' and 'arrestment' of an exhibition to a visitor were defined in terms of the number of exhibits he stops at and his average (exhibit) viewing time, respectively. It was postulated that for a visitor, the 'attractiveness' and 'arrestment' of the exhibition would be related to his or her 'prior interest' in the topics. Displayed prior interest was defined as the interest a visitor brings with him to the exhibition and it refers to his prospective interest in the subject matter he will encounter on display quite independently of his actual interest in the exhibit as he eventually finds it.

A sample of 148 visitors was interviewed before it entered the Hall of Human Biology to obtain measures of prior interest in the topics on display; and the behaviour of subjects in the exhibition was observed unobtrusively using the apparatus previously described in Chapter 4. The attracting powers and holding powers of exhibits and Assemblies (see Chapter 1) were computed and reported in detail along with summaries of all the time data and measures of prior interest. By and large, the time data have replicated the findings of earlier studies, indicating that on the average visitors spend approximately twenty seconds looking at individual exhibits. This provided empirical evidence to add to the philosophical considerations with which to question the usefulness of 'holding power' as an index of exhibit effectiveness. It seems that visitors are simply not prepared to spend more than thirty or so seconds looking at individual exhibits.

Although replicating certain of the findings of earlier researches, some of the other effects reported in the literature were not apparent from the study. The pronounced 'right-turning' tendency as shown in all studies in American museums was not found; the

so-called 'exit-gradient', whereby the attracting power and holding power of exhibits are related to their physical distance from the exit door was also missing; and 'museum fatigue' which Melton showed to affect attracting power while leaving holding power unaffected was not noticeable.

There was a significant positive correlation between an individual's prior interest in the Human Biology exhibition and the attractiveness of the exhibition but no relationship between prior interest and the arrestment of the exhibition. In other words, these findings suggest that individuals are likely to stop at more exhibits in exhibitions dealing with topics of interest to them than they are in exhibitions dealing with topics of little or no interest to them. Knowing prior interest removes some of the apparent randomness from visitor behaviour.

The failure to find any correlation between prior interest and arrestment furnishes yet further empirical evidence that visitors are only prepared to spend a limited amount of time looking at individual exhibits; and attempts by designers to capture visitor's attention for longer periods at single exhibits are likely to prove unfruitful. Nevertheless, in absolute terms, the correlation between prior interest in an exhibition and its attractiveness was low ($r = 0.25$), thus only 6% of the variability in visitor behaviour was explainable from knowing prior interest. This clearly indicated that there were other factors or perceived characteristics about exhibits which caused a visitor to stop.

Two studies were reported in Chapter 6 pursuing this theme. The first was concerned with elucidating the characteristics of exhibits in the Hall of Human Biology as perceived by visitors. These characteristics were elicited from a sample of visitors taken to exhibits in the exhibition using a free-response questioning format.

The large number of responses was reduced to a representative list of approximately fifty items. This list was used by subjects in the second study who were invited to state which of the items on the list applied to a selected list of exhibits. Forty five exhibits in the Hall of Human Biology were 'rated' in this way. The exhibits were chosen to be representative of the full range of 'morphological types' to be found in the exhibition; and they were also representative of the ranges of attracting power and holding power calculated in the studies reported in Chapter 6. In addition, subjects were also asked to state which of the items would apply to an ideal exhibit. Using a Biplot analysis (a sort of principal components analysis for rows and columns simultaneously) it was possible to see which items applied to which exhibits and which exhibits were similar to each other in terms of the relative attributions of the items to the exhibits. Moreover, the distances of each real-world exhibit from the ideal were computed and the values obtained were correlated with the exhibits' attracting powers and holding powers. Characteristics which applied strongly to the ideal were:-

- it makes the subject come to life;
- it gets the message across quickly;
- you can understand the point/s it is making quickly;
- there's something in it for all ages;
- you can't help noticing it.

Characteristics which were strongly negative with respect to the ideal were:-

- it's badly placed - you wouldn't notice it easily;
- it doesn't give enough information;
- it's not clear what you're supposed to do or how you should begin;

- your attention is distracted from it by other displays;
- it's a difficult subject to understand.

Characteristics orthogonal to the ideal were:-

- it allows you to test yourself to see if you are right;
- it involves you;
- it deals with a subject better than textbooks do;
- the information is clearly presented;
- it makes a difficult subject easier.

The correlation coefficient between the distances of exhibits from the ideal exhibit and their attracting powers was $r = -0.63$, highly significant at $p = 0.01$, whereas there was no significant correlation between exhibits' distances from the ideal and their holding power ($r = -0.09$).

Interpretations of these results is favourable to the theory of exhibit effectiveness (outlined in chapter 3) in which real world exhibits were conceptualised in terms of their proximity to an ideal exhibit. The proposition that exhibits are in competition with one another with successful ones being those that attract visitors' attention is given convincing support by the evidence. Explanatory power is added to the theory on the basis of the results by showing that an attractive exhibit is one that transmits a short clear message, displayed in a vivid manner with a wide appeal to all age groups. These findings appear to relegate some of the more ambitious interactive exhibits to the status of ineffective exhibits, particularly those that require the visitor to invest a considerable amount of his or her time reading text and instructions.

The strictly behavioural data collected in the studies reported in Chapter 6 called into question holding power as an index of

of effectiveness by showing that visitors on average spend only a short time looking at exhibits. The studies reported in Chapter 7 have provided yet further evidence to this effect by showing that the holding power of an exhibit is unrelated to its distance from the ideal. This is corroborated by the attributes visitors associate with the ideal exhibit; specifically, that "it gets the message across quickly" and "you can understand the point/s it is making quickly."

Weiss and Bourtourline (1963) speculated that unusually large or vivid exhibits, which they named 'landmark' exhibits, would attract visitors and, therefore, be determinants of paths through the museum. The three exhibits closest to the ideal - 3, 38 and 26 (Figure 17) clearly support the idea of landmark exhibits. Exhibit number 3, consists of four short slide-tape presentations in which the visitor is confronted with dramatic depictions of a series of familiar happenings that give rise to the emergency reaction. Here, sound is the most apparent form of initial attraction. It announces the presence of the exhibit well before it can be seen, and a frightening scream which is part of the information - bearing elements of one of the four sequences, is also effective in attracting attention. Initial attraction, however, although necessary, has to be reinforced if the visitor is to be held long enough to receive information. In this exhibit, reinforcement is provided at the outset by a question incorporated in the first unit of the section ("Emergency - How do you react?"). Since the design of the exhibit incorporates slides with accompanying sounds, timing becomes an important factor. It has been shown, for example, that most of the relevant information in a picture is extracted during the first two seconds of eye fixation (Mackworth and Morandi, 1967); and since the slides are clear and the main focus is emphasised, it might be proposed that the rapid cut from slide to slide which is a feature of the

exhibit, manages to retain the attention of the viewer - the median holding power of this exhibit was 55 seconds which indicates that at least fifty per cent of the sample remained long enough to receive the main points displayed.

Exhibit number 38, known as the Womb Area, consists of a separate darkened room containing a back-lit, giant-sized model of a human foetus, accompanied by sound recordings taken from a mother's uterus. Size and location are two forms of attraction particularly exploited by typography, although there is evidence that these are significant pictorial factors as well (Brandt, 1945); and the large foetus in the Womb Area is a good example translated into three dimensions. However, there are three other powerful elements which might also account for the exhibit's attractiveness. There is what might be called the 'immediate meaning' of the exhibit, to the extent that if the viewer were unaware of the identity of the object, its particular power to attract would be diminished. The most well-documented evidence of universal interest (for example see Buswell 1935, and Yarbus 1967) concerns the compelling attraction of human faces; and it is perhaps reasonable to extend these findings to the three-dimensional model of a human foetus - the beginnings of life might be safely assumed to be of interest to most people. Finally, there is isolation, not of the exhibit but that of the viewer himself. The visitor to the Womb Area is not in an ordinary room but in an environment in which authentic sounds, acoustic design, and comparative darkness combine to provide a veridical set of identifying features approximating to a womb.

Exhibit number 26 is a large exhibit in comparison to others in the assembly and it is centrally placed, thus being in a sense difficult to ignore. The exhibit consists of a series of models which are manipulated by the visitor, showing different joints found in the

human skeleton and how they work. It is 'visibly' interactive and this seems also to add to its attractiveness. In this exhibit, the quality of the interaction is dictated by the nature of the information and may therefore give the exhibit an 'immediate meaning' to the visitor.

It should be stressed that although all three of these exhibits are well described by the statement "You can't help noticing it" (closely associated with the ideal), this statement alone is insufficient to characterise the exhibits' close proximity to the ideal. For example, Exhibit number 19 is a model of a generalised human cell, situated at the beginning of the Human Biology exhibition. It is literally the first exhibit on view to visitors as they enter the exhibition and it is large; and yet, relatively speaking, it is some distance from the ideal. As well as its size, it is to some extent physically isolated; further, the lighting and colouring of the three-dimensional part of the exhibit gives it both tonal contrast (Buswell 1935) and colour contrast (Chevreul 1967) with a dark blue background. The information presented is designed to show the complexity of the cell. There is a cross-section in the model that presents a visual array containing more than two dozen elements, differing from each other to some degree in shape and/or colour. The textual information is accompanied by colour photographs (Dwyer 1971) with coloured headings (Brandt 1945) and enough white space within and around the photographs to suggest easy access (Spencer 1969). Accordingly, one might have expected this exhibit to have turned out to be very attractive since it conforms with many of the principles of good design for communication and it is well located. It seems that the simple answer as to why it is not particularly attractive is that it takes the visitor too long a time to assimilate and understand the information presented; and this fact is evidently signaled to visitors before they stop at the exhibit. It is as if

visitors perceive the threat of suffocation if they dare to stop. This suggests that one of the factors which attracts visitors to exhibits is the (perceived) pay-off between stopping and receiving information without the necessity of having to invest too much time and effort. The question of whether it is possible to induce a visitor to stop at an exhibit in which a more significant investment of his time is required to receive the exhibit's message will be discussed later in the chapter. However, at present, it seems that effective exhibits are those that attract the visitor and then 'pace' the presentation of a limited amount of information. If the visitor perceives the information to be complex he is unlikely to stop. Effective exhibits are thus those that have an 'immediate meaning' to the visitor.

Similar points to those made about exhibit 26 could be made about many of the exhibits some distance from the ideal. They all pre-require from the visitor an investment of time he or she is apparently not prepared to make. Some of these exhibits are the ones that the planners and designers of the Hall of Human Biology have singled out for special mention as being examples of the sorts of 'ideal' exhibits that museums should produce. Miles and Tout (1978) referred to exhibits in section J ('Cogdevel') of the exhibition which present the visitor with a series of logical puzzles to solve. These are all exhibits which require the visitor to read detailed instructions on how to interact with them before they can even begin to understand the message. For interactive exhibits to be successful it seems that the act of interacting must itself be informative and that interaction as a means to obtain information is not good enough.

These conclusions have rather profound implications for designers of educational exhibitions who have assumed that museums have a

potential for learning in much the same way as textbooks and programmed learning material. This formal educational model has been shown up as being clearly inappropriate at least as presently interpreted in the museum setting by the research presented in this thesis.

As to what alternative approach might prove satisfactory is, of course, an altogether different question. Perhaps one of the first things to be done is to remove the wholly artificial dichotomy between museums as places of entertainment and museums as learning environments. Visitors came to museums first and foremost to enjoy themselves as countless visitor surveys show (see Appendix 1); and it is the 'experiential' aspects of a museum visit that should receive attention from museum planners, particularly with a view to designing exhibits which have an immediate meaning to them. This would entail moving away from thematic exhibitions with rigid learning hierarchies to exhibitions with networks of inter-related ideas and objects which different visitors could explore in all manner of different ways as their own interests dictate.

Before turning finally to a discussion of the possible directions future research would take in this area, the relevance of the present approach to wider fields of psychological study will be discussed briefly.

As mentioned in Chapter 3, the study of 'curiosity' could be illuminated by a general theory of attractiveness couched in somewhat similar terms to those used in the more specific theory of effective exhibits. If curiosity is considered as a 'state' aroused by certain objects, the methodology outlined in Chapter 7 could easily be applied to the study of individuals' perceptions of different objects; and in doing so it would provide a more complete

explanation of the curiosity-value of the objects than ones offered in terms of the stimulus complexity of the objects themselves. Further, the idea of curiosity as a 'trait' could be explored by examining individual differences in the 'attractiveness' of an object, perhaps in terms of relevant personality variables. These observations also apply more generally to the study of aesthetics and studies of liking or preferences for objects.

Perhaps also, the studies reported in Chapter 7 have a relevance to the wider field of cognitive psychology. Rosch (1978) has made a plea for more naturalistic research into the manner in which we categorise real-world objects since most of the studies reported have been undertaken out of context in artificial laboratory conditions. It has been demonstrated in Chapter 7 that it is possible to use sound research - and even experimental - procedures in a natural setting. These procedures have yielded valid results in the sense that they correlate highly with a behavioural index of exhibit effectiveness computed from data collected from unobtrusive observations of museum visitors in a separate study. Furthermore, the correlations obtained are at least as high as might be expected from studies carried out in a controlled laboratory environment. In addition, the methods reported suggest themselves as a general method for investigating the ways in which we perceive similarities between real-world objects.

Further research directions

A number of avenues for further research are suggested by the findings presented and discussed in this thesis; and some of these will now be presented.

A Trade-off Model for exhibit characteristics

As pointed out in Chapter 7, certain of the characteristics which turned out to be unrelated to the ideal ran counter to intuition, although some of these findings may be partly explained by the overall structure in the data. Nevertheless, the salience of various characteristics does merit further investigation, in particular to see if and under what circumstances characteristics closely associated with the ideal might be traded-off against attributes which are less favourable or unrelated. For example, the characteristic "It involves you" was found to be orthogonal to the ideal exhibit but as pointed out, this should not be taken to imply necessarily that an ideal exhibit should not possess this characteristic. If an exhibit were perceived to possess the characteristics associated with the ideal and "it involved" the visitor, this may be highly desirable. The question then arises under what circumstances would visitors trade-off one or more ideal characteristic in exchange for "It involves you". As an illustration, consider just one characteristic associated with the ideal - "It gets the message across quickly" - which visitors might trade-off against "It involves you." In principle, it is possible to specify various levels of involvement and various required viewing times for a hypothetical exhibit. If each characteristic is specified at three levels, then subjects could be asked to rank all nine combinations of involvement and required viewing time, thus discovering the extent to which a subject would be prepared to trade-off required viewing time against some level of involvement. In this illustration, only two characteristics have been suggested but it is possible to use trade-off computations between combinations of the levels of any number of attributes (Kruskal 1965). This model would indicate which hypothetical exhibits would be preferred to which other hypothetical exhibits, thereby demonstrating possible strategies for a designer

to adopt in a given context. Such an approach would indicate, for example, under what circumstances if any, visitors might be prepared to view exhibits for longer than the apparently immutable thirty or so seconds.

Validating and Extending the Paradigm

While the above line of research is a natural consequence arising out of the present research programme, the immediate thrust of any future research should concentrate on validating and extending the paradigm established for testing the theory of exhibit effectiveness outlined in Chapter 3.

The research reported in this thesis was carried out in a single exhibition in the Natural History Museum and should be tested further in other exhibitions employing different approaches to design and in different museums covering different subject matters. Clearly, to borrow Repertory Grid terminology, the range of convenience of the characteristics will vary from exhibition to exhibition not to mention from museum to museum, and this means that relevant and appropriate characteristics will need to be elicited for each study. A failure to do so would possibly subvert the subsequent stages in the research paradigm. However, while it is obvious that the characteristics associated with real exhibits will vary in different exhibitions using different styles of presentations, it is by no means as certain that the characteristics associated with ideal exhibits will vary as greatly. It is an empirical exercise to discover if visitors have different ideals for different exhibitions.

Following on from these considerations, there is a need to refine operationally the notion of an 'ideal' exhibit to take account of individual differences among visitors. As discussed at the end of

Chapter 7, the present research self-consciously took a consensus across the sample to arrive at a profile of the ideal exhibit. As it turned out, this was a reasonable stand-point to have adopted but it would have been much more sensitive to have looked for individual differences in ideal exhibits since by doing so firmer predictions and greater explanatory power would be anticipated as a result. However, exhibition planners and designers are usually going to be concerned with populations of visitors rather than with individuals and there is thus a limit to the extent to which the population of all visitors can be disaggregated to take account of individual differences. One approach to segmenting the population into sub-populations would be to investigate if, within the total population of visitors, it is possible to discern different 'clusters' of individuals with identifiably different requirements for their ideal exhibit. If noticeable sub-populations were detected, then it would be possible to attempt a definition of these sub-populations (possibly in terms of their demographic characteristics), thereby defining target audiences for different approaches to exhibition design.

One of the hoary old chestnuts in the museum world centres on the perennial question "What is a museum?" Museum professional pontificate endlessly (and fruitlessly) on this question since the question can never be illuminated by recourse to reason alone. A museum can be anything (within reason) that its managers or trustees decide it shall be. However, the question does take on a different complexion if posed in the form "What do visitors think a museum should be?" Using the research paradigm developed during this thesis, it would be possible to elucidate how different museums are perceived as institutions and relate these perceptions to the notion of an ideal museum. It may be that different 'ideals' exist for different types of museums - science museums and art galleries, for example, may be

characterised by different ideals. At least in this way the question becomes tractable; and in the age of accountability, museums could begin to take notice of their public.

Finally, and in deference to those colleagues preoccupied with the idea of the exhibition as an educational medium, a corollary of the findings presented in this thesis is that visitors should learn more from exhibits close to the ideal than those distant from it. This is a relatively straightforward hypothesis to test but the present writer's experience in evaluating the Hall of Human Biology over the past five years, suggests it would miss the point concerning what exhibitions are about.

Appendix I

"Four years of visitor surveys at the British
Museum (Natural History): 1976 - 1979."

Text cut off in
original

Appendix II

Descriptions presented to subjects for ratings
of prior interest.

(see Chapter 6)

A. Living cells.

This section of the exhibition explains how the different cells in our bodies carry out different jobs. The sperm cell in the man and the ovum in the woman are discussed in more detail and, in particular, how these two cells join together to form the first cell of a new Human Being.

B. Growth and development.

This section of the exhibition traces the growth of a baby in the mother's womb up to the time of birth. It explains how a baby is born and how a baby grows during childhood.

BE. More about chromosomes.

This section of the exhibition explains how we inherit genetic instructions from our parents.

C. Movement.

This section of the exhibition explains how muscles pull our bones to make us move. It explains how different joints allow our bones to move in different ways, how muscles pull our bones and how nerve messages control the movements of our muscles.

D & DE. Controlling your actions.

This section of the exhibition explains what the brain - the body's control centre - is made of and it explains the different areas of the brain and how they work together to control your actions.

E. Your life in the balance.

This section of the exhibition explains how conditions inside your body stay more or less the same despite the fact that our surroundings are constantly changing.

F. Hormones - messengers in the blood.

This section of the exhibition explains how hormones control different bodily processes. It also explains what hormones are, how they work and how they themselves are controlled.

FE. More about sex hormones.

This section of the exhibition explains how hormones control the development of our sex organs. It also explains how sex hormones control a woman's monthly periods.

G. Hormones and nerves.

This section of the exhibition explains how hormones and nerves work together to control the amount of water in our bodies.

H & HE. Learning and memory.

This section of the exhibition explains how we learn skills and information, how we remember the things we have learned and the methods we use to recall information from our memories.

I. Perception.

This section of the exhibition is concerned mainly with how we

perceive objects in the world about us. It shows through the study of visual illusions how the brain interprets information received through our eyes.

J. How do we come to understand the world about us.

This section of the exhibition traces the stages of intellectual development we pass through from early infancy to adulthood. It shows the importance of play and imitation, traces the development of language and our ability to solve difficult intellectual problems.

Appendix III

Photographs of exhibits studied (see Chapter 7)

Notes:

- (i) The photographs of the exhibits are numbered as in Table 18.
- (ii) Unfortunately photographs of exhibits 13, 17, 19, 21, 34 and 41 are missing as these exhibits underwent revisions shortly after the data were collected and before the photographs were taken.

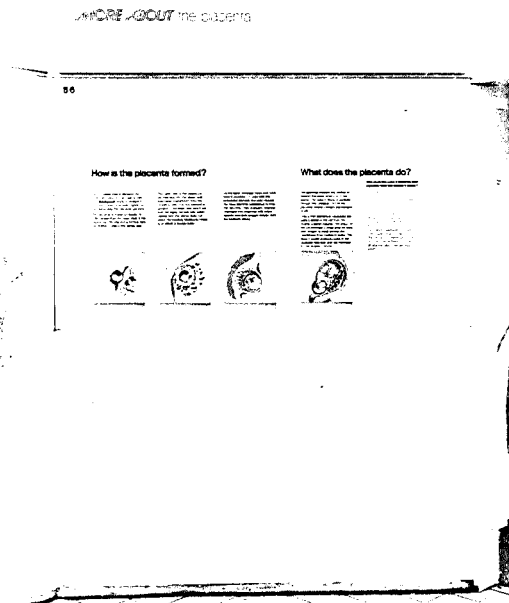


Exhibit 1.

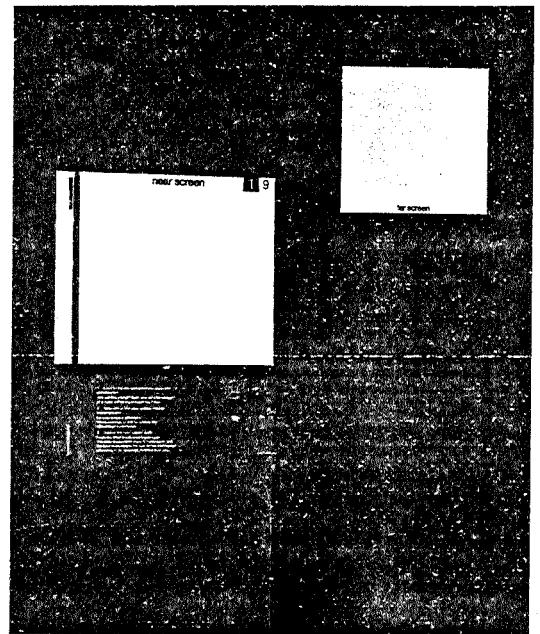


Exhibit 2.

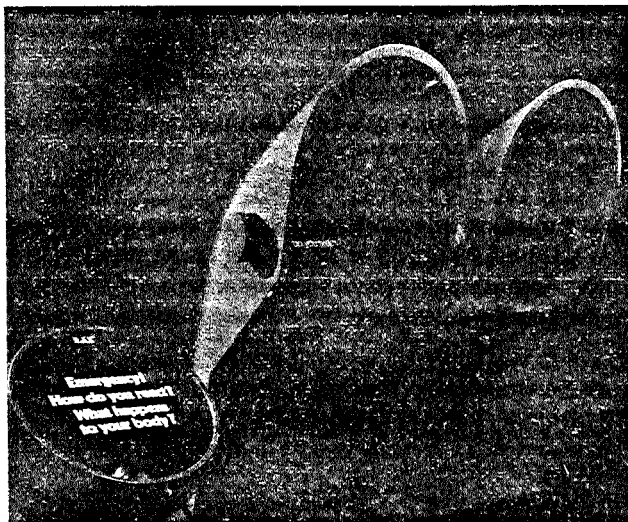


Exhibit 3.

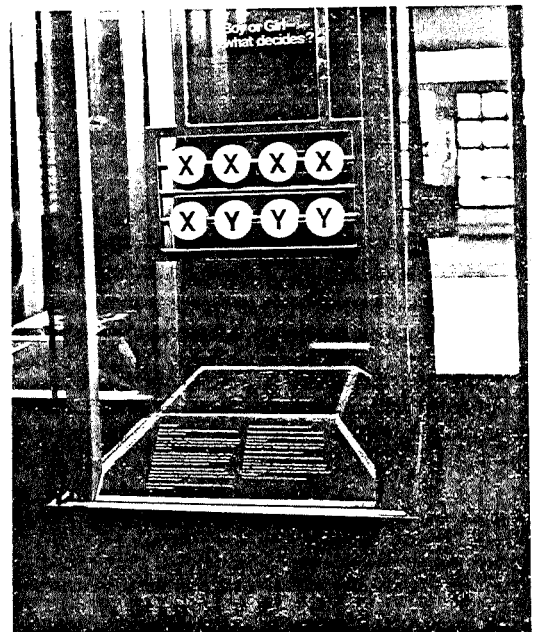


Exhibit 4.

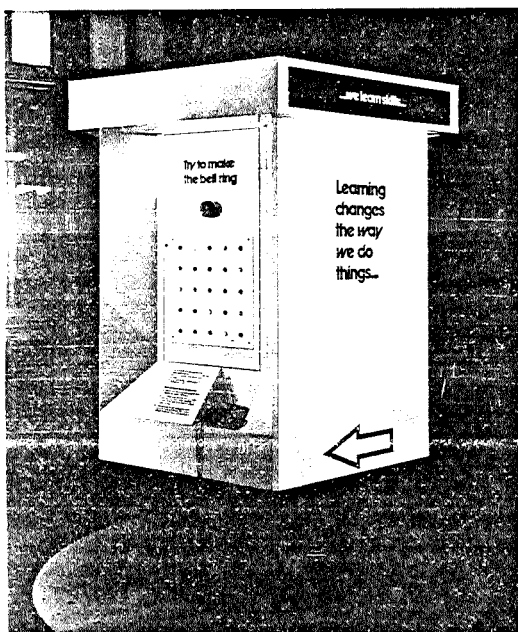


Exhibit 5.

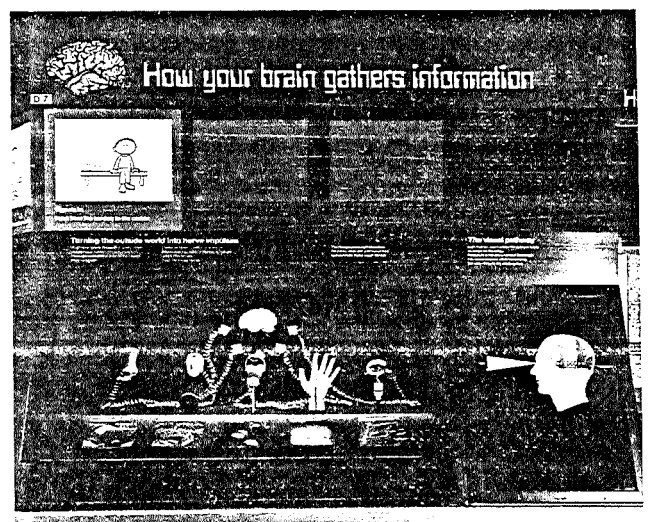


Exhibit 6.

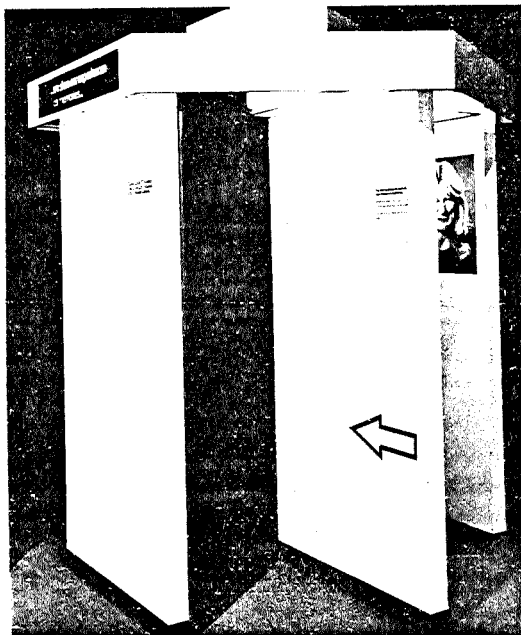


Exhibit 7.

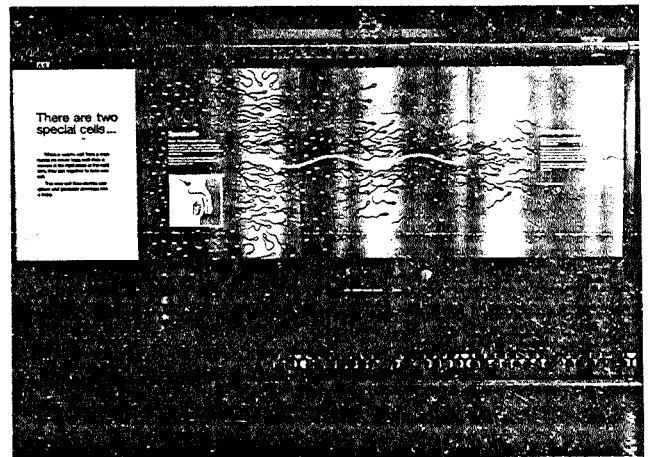


Exhibit 8.

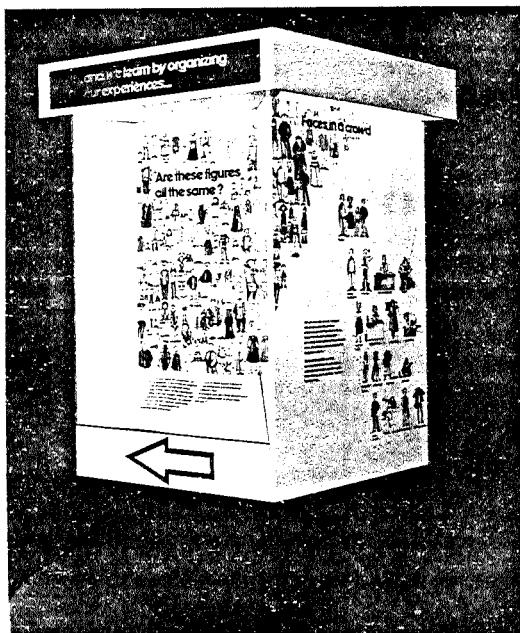


Exhibit 9.

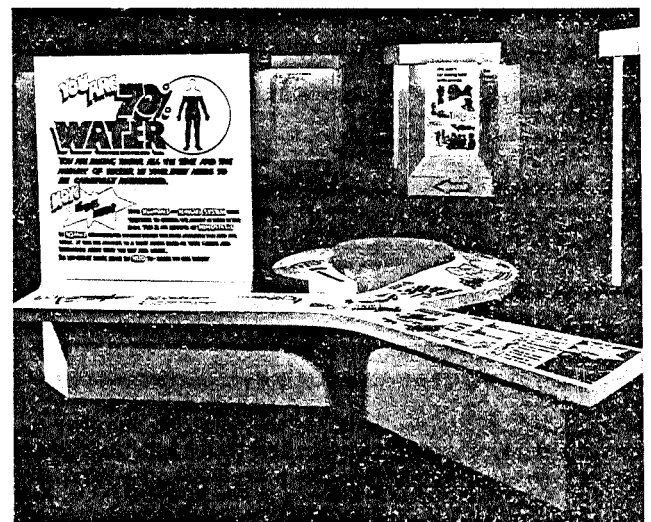


Exhibit 10.

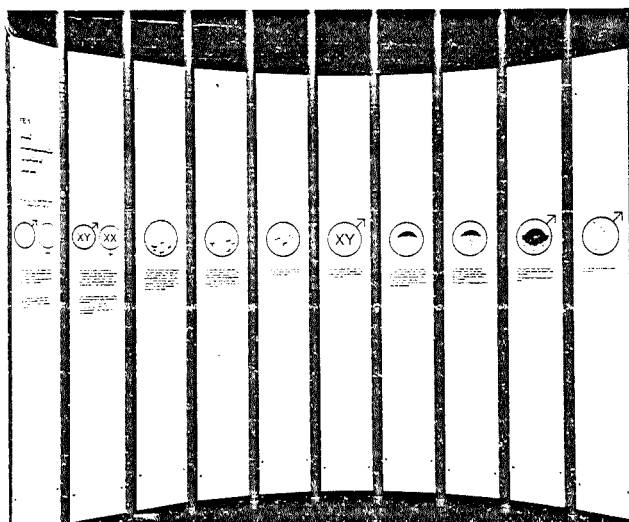


Exhibit 11.

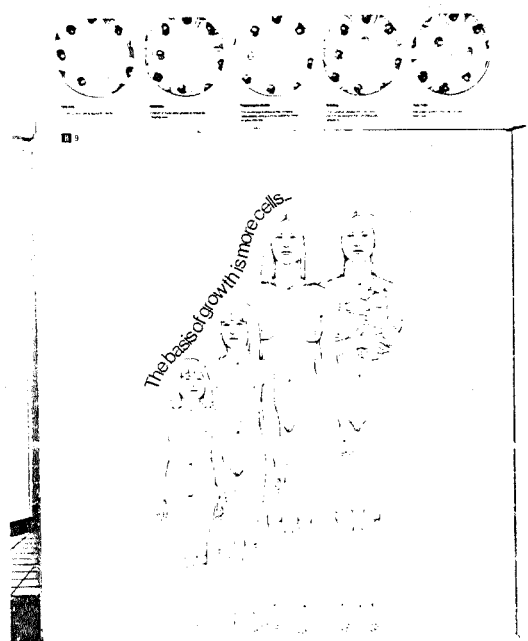


Exhibit 12.

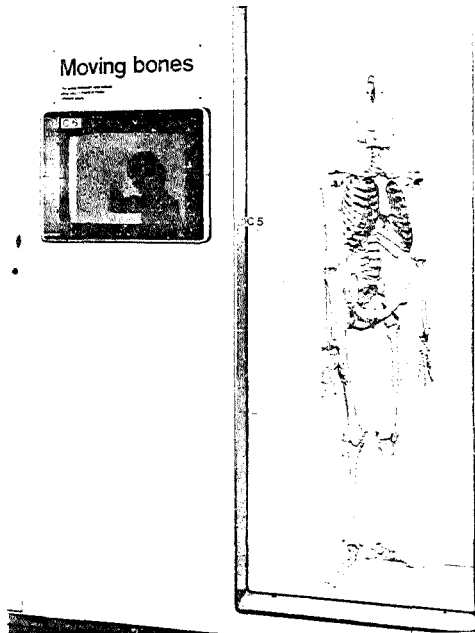


Exhibit 14.

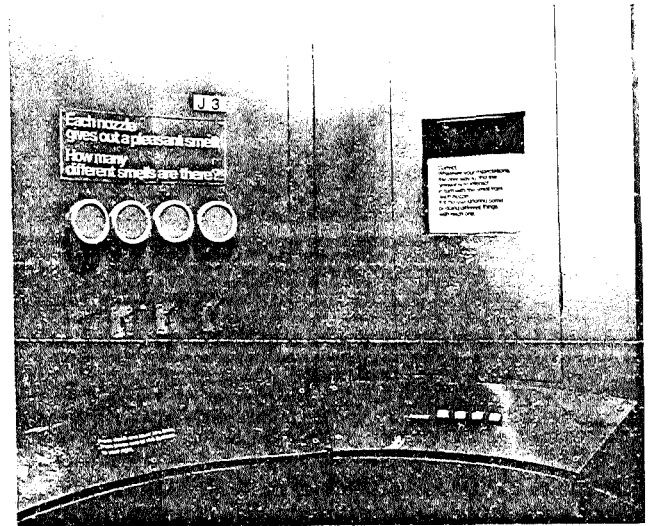


Exhibit 15.

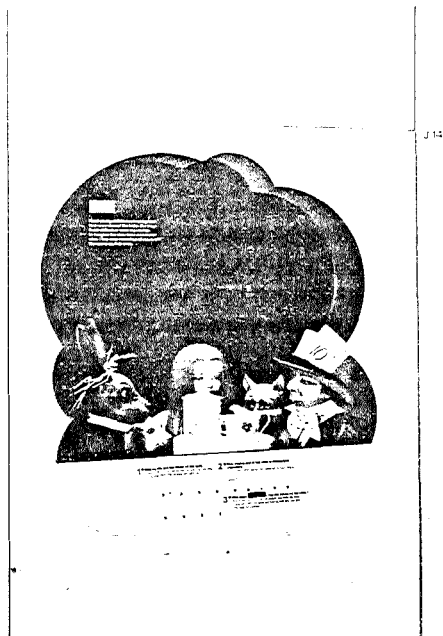


Exhibit 16.

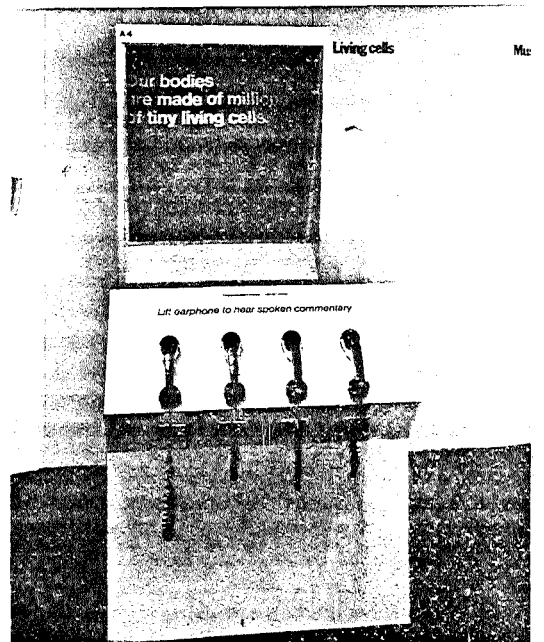


Exhibit 18.

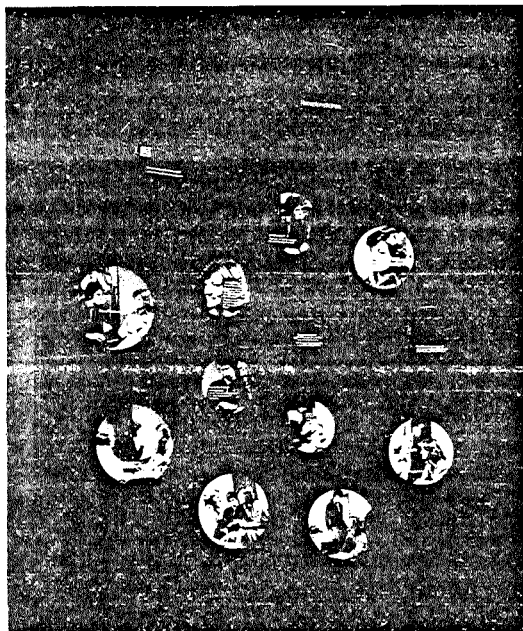


Exhibit 20.

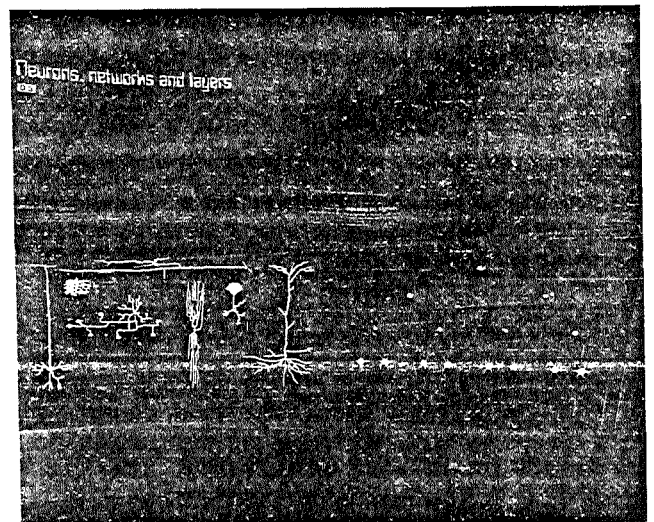


Exhibit 22

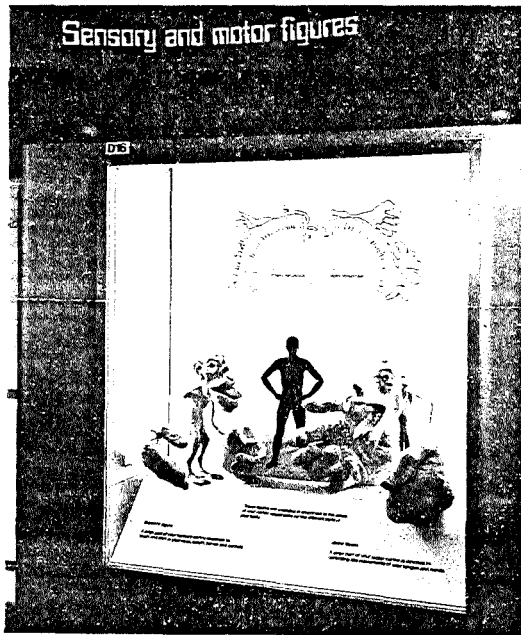


Exhibit 23.



Exhibit 24.

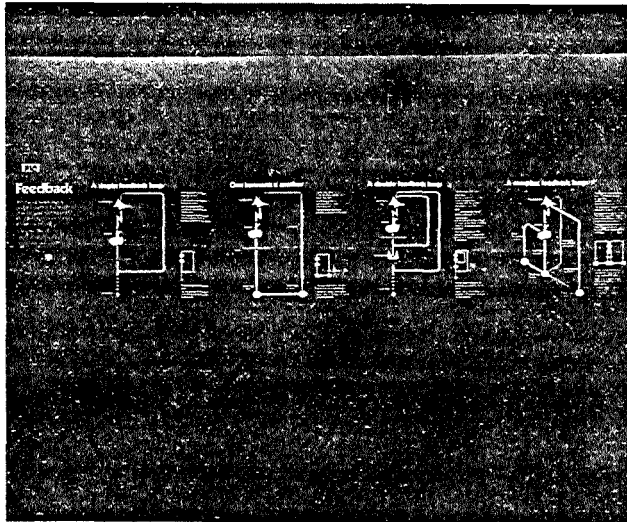


Exhibit 25.

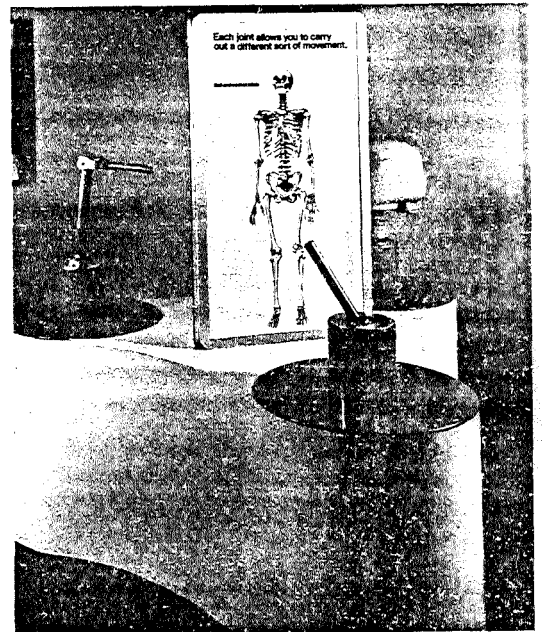


Exhibit 26.

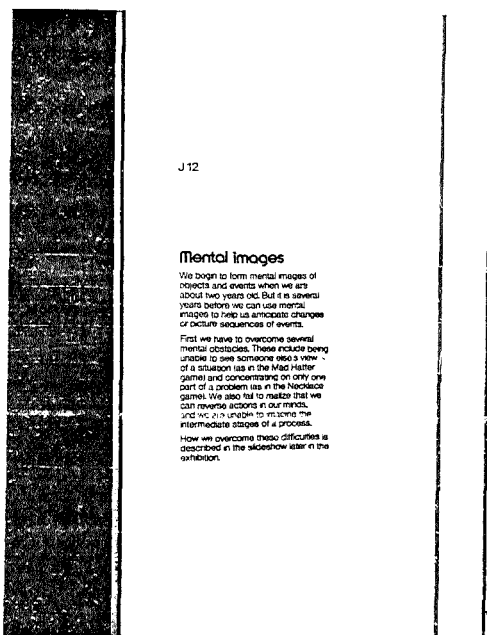


Exhibit 27.

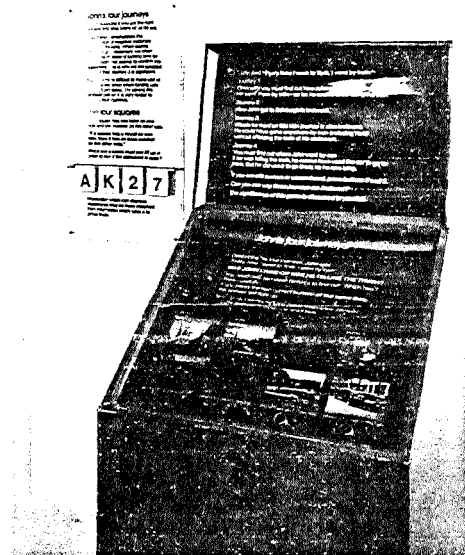


Exhibit 28.

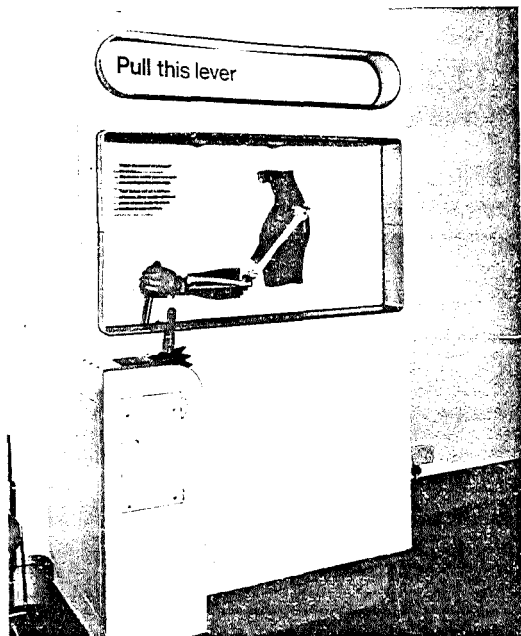


Exhibit 29.

Can you add a third ring and still
make just one shadow?
What controls the size of the
shadow?
A. C.
The number of rings on the stick.
The size?
The colour of the rings?
What the rings are made of?
The distance from the light to
the rings?
The answer is on the wall beside
the game.

225

A shadow is produced
by an object blocking the light
from a source. The size of the
shadow depends on the size of the
object, the distance from the
light to the object, and the
distance from the object to the
screen.

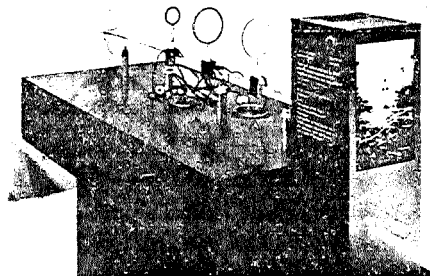


Exhibit 30.

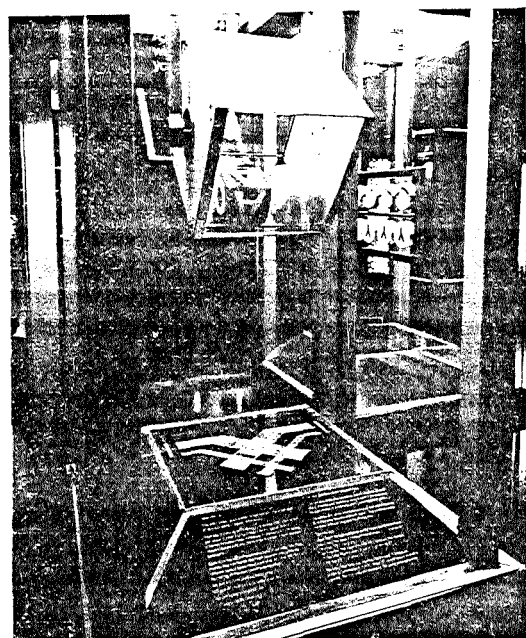
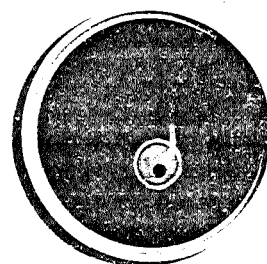


Exhibit 31.

A 10

How the special cells meet

A few hundred sperm cells meet
the egg. The sperm, but only one will
fertilise it.



Two millions in the human ovum
and two millions of spermatozoa
swim from the ovule and the sperm
cell to the egg.
In this special case to your left, you
can see that only one is important
in the union to your right, you can
see that the sperm cell is the only
one that is a baby.

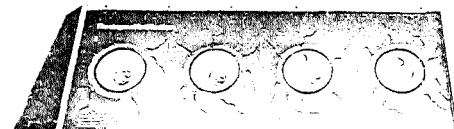


Exhibit 32.

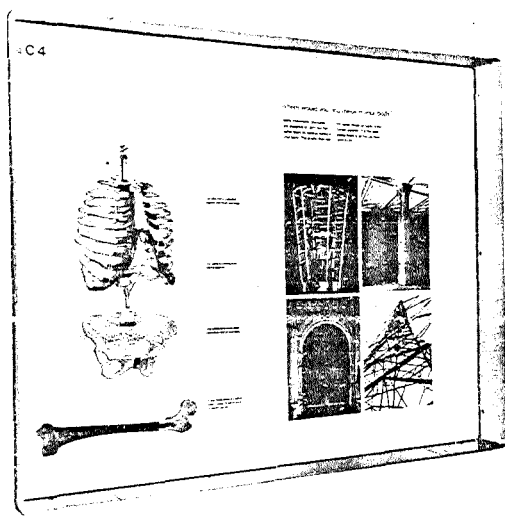


Exhibit 33.

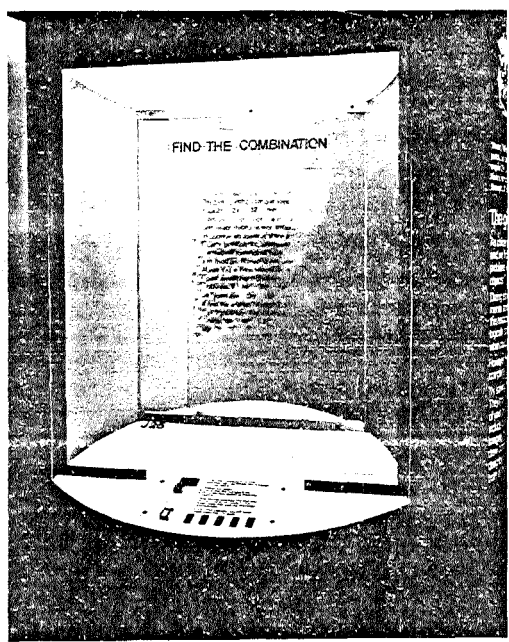


Exhibit 35.

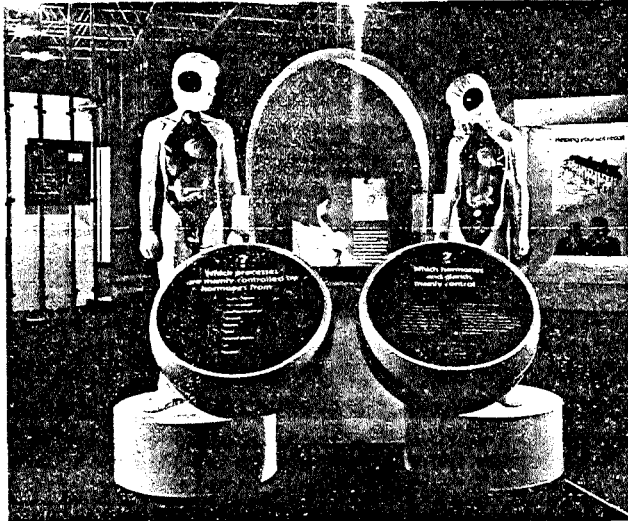


Exhibit 36.

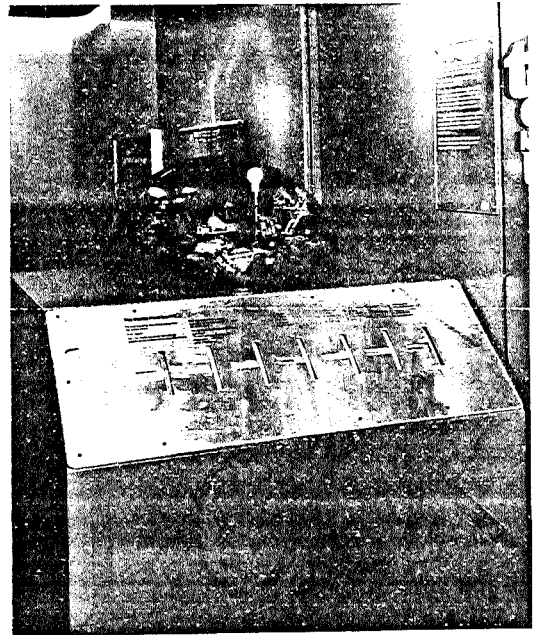


Exhibit 37.



Exhibit 38.

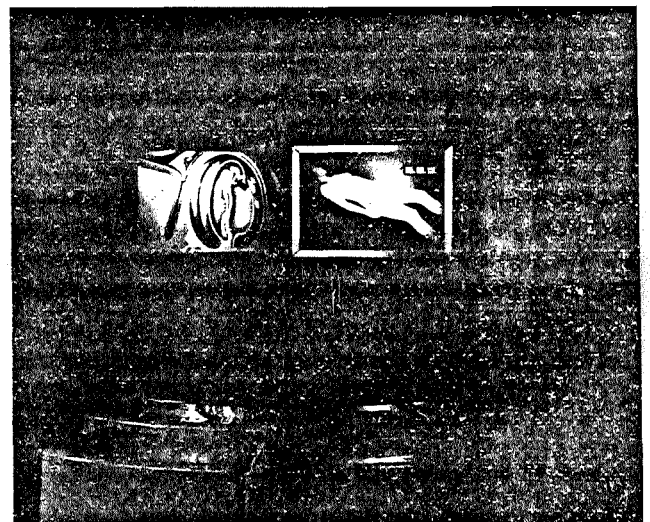


Exhibit 39.

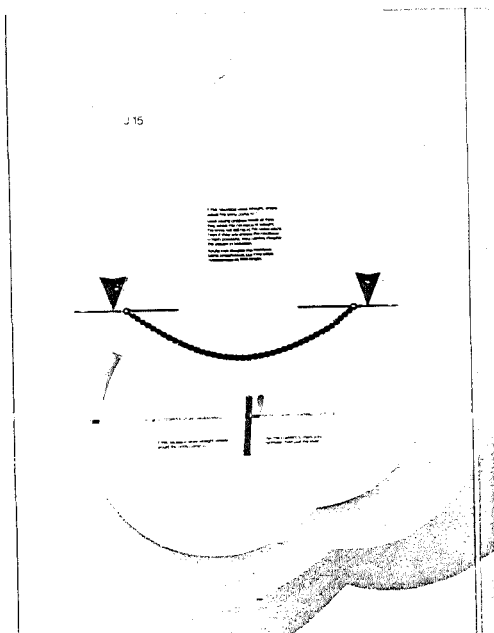


Exhibit 40.

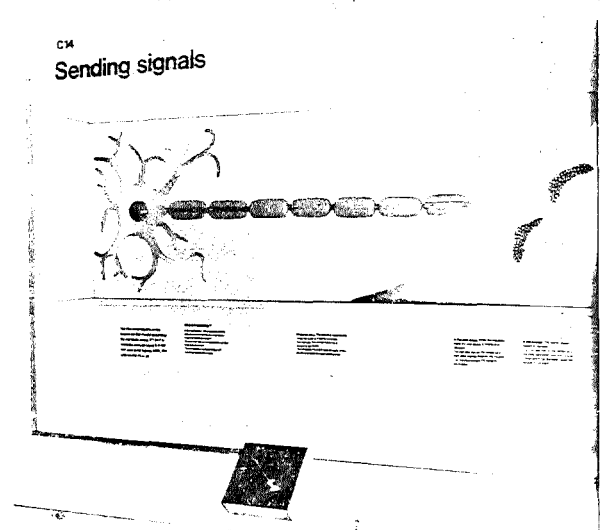


Exhibit 42.

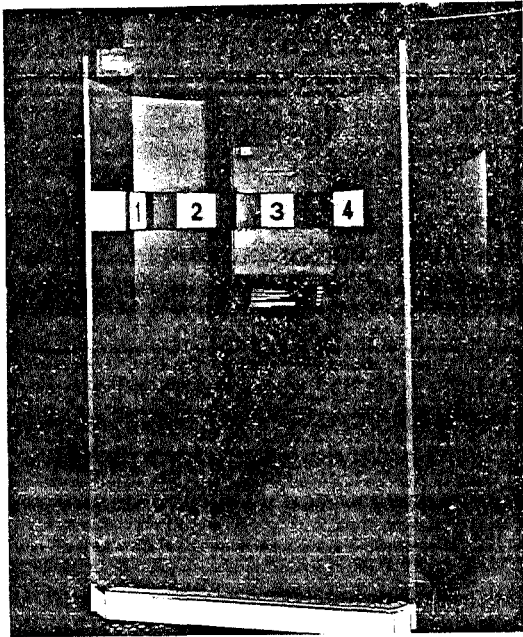


Exhibit 40.

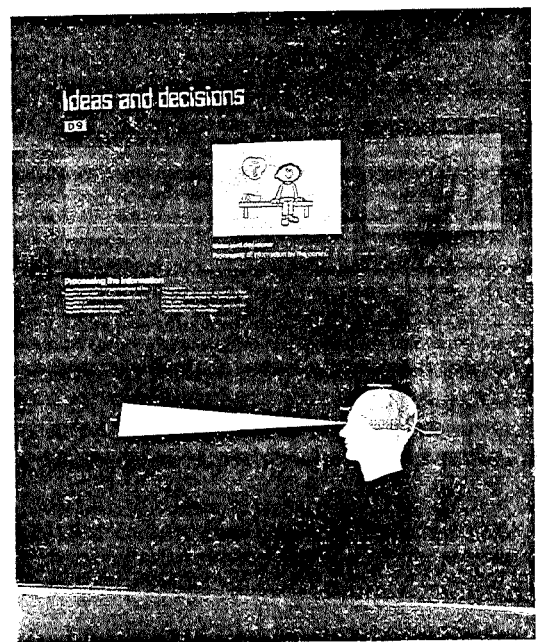


Exhibit 44.

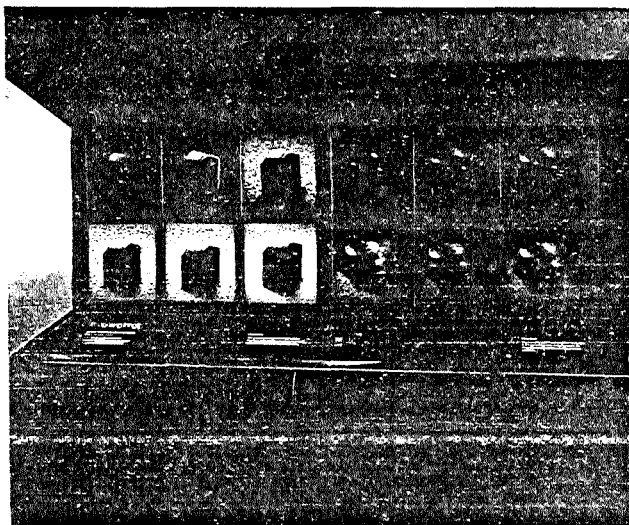


Exhibit 45.

Appendix IV

Complete percentage matrix used as input for the biplot analysis (see Chapter 7).

Note:

- (i) Columns represent exhibits and run in sequence from exhibit 1 to exhibit 45 (Table 18) with the final column representing the ideal.
- (ii) Rows represent characteristics and run in sequence from item 1 to item 53 (Table 20).

USER DISK JOB FX YY/MM/DD TIME--ON TIME--NOW CPU-TIME PAGE 1
KM2 01 13 RC/13/31 15.35.29 15.55.59 M:38:395 NP/3 NK 6.1

33 3971737732784 86213341729 56336348350332666233072 0 960172354 82236745590563819665145 762
20 42020 8 3 1 228395714 7 0 2 91282 852 2 7 2 048 11024 01930 7 6 0202019 274 0 0 3 1 2 3 9
1222631616752 138 7 010 717423842 21811 2104222 732 033484034 0 21934 845 8 037203046 52615
021 1 9 6 330 114 6 0 2404813 323 2 120 123 122 41027151210 8 12652 5 117 2 0 821 6131116 0
253 032743462 169 2 6 018 0827722 1 2 6 115 324 320 174426928 2 3197451 9 3 08651 568 38471
1916 646 4 0 01316 93211 2 1 6 3 832 831 63116 056 01636 2 94718 8 34317 7 3 9 3 416 313 2 8
2617 259 9 1 1 82716392011 0 51124232514225620 174 1 831 01567 7 6102115 9 2 9 2 834 419 213
0 2 6 3 41430 2 1 1 0 133 9 5 4 1 6 024 1 1 124 01910 2 9 7 0 01727 0 0 8 5 1 9 8 315 315 0
929 03716 2 6 34434 01415 0162740 5 714 11611 032 03333 14445 2 61610 752 0 0 8 414 110 0 3
0 7 2 930 321 220 1 0 011 11521 7 0 0 1 0 2 0 2 1 2 128 019 3 0 0 737 214 0 018 3 211 211 0
0 0 3 1 1 2 0 8 0 1 0 216 2 1 0 7 142 2 22645 7 6 1 0 0 0 1 1 1 0 5 0 36756 0 0 728 013 121
89385155508835933272948220143560517591669131647262676328763848917837348520729771657367775994
3220 2 7 9 3443419243632526716 4 31035172830491716 742 7 0 0 5 1495315145943 6 027113713 9 1
5019314614101030272847381616101528364125 8624711683620302023494525131753172218 61939 9322631
2712 714 3 8 310 3213315 4 514 9 81215151712 7 312 9 2 3 7 2 713 0 5 226 62640 2 9 8 516 335
54787 6816540 64726 11110 261813628 734 7 33976 643 22056361428 3 8695347335451473060 95759
140 930221247 533 5 110617736 53813 533 758201712313439212423 7316316 23811 0102719183622 0
45 115 5 117 758 841253316 1 030 12443132523261039 1 3 2 3 1 823 717 027185127 2 339 1311740
93078 9387964 74054 6383315726740 7 07086 52023 439 248442227 27617 611 515196334 547177566
4225352813442261 523272826451323132020501121455194531 3332621524328 522 54854 94920333021 2
021 03410 110 04813 2 521 117 717 0 5 5 245 5 030 0 642 03036 1 21527 033 0 0 3 2 6 9 3 6 0
284277 8715356493443 841273356733531581370105550 868 3 9763616513527295645867978703973406784
131671 62314 8271911 911 4 517481610732553 864161111 3 416 2 820 3 32158139057 612271022 543
15753138885879178013111122 690777039 6181121 7151068 773888344 9 542915977502586843387117979
7022 8671421 4393141766312 8 92640587227525138 373162633 919724418 92168142765 61665 9601049
35256 6907238114521 110 7 6777736 8 513 2 63400 338 117652316 6 5 870383923 861541167 35252
191037171363 656 7131333 9 9 247 8 832613836402458 6 2 43113243211 2 85733411123 95925681551
2 1 1 0 85017 5 1 1 0 61216 1 2 8 1 0 2 1 2 156 016 0 020 5 0 13010 0 1 8 6 011 4 029 5 7 1
342262572939425346304019343669533061254030612553361242304253201130596351353847345339383787
130 13210 528 04111 2 95716381420 2 219 053 5 0 7 11446 14036 3 73817 856 2 0 0 4 9 4 7 7 0
612736280492286417487743 722283323486834811961154863 3 95130357234 81580 6779728536134512455
4421 311 7 1 2 918444914 6 01413202352 5631543 13811 3 4 2 917 611 6 759 6194447 7231522 660
3927 9262913123228385344 9 5192341263438132939 031142933 73833261717373431 628 91236133415 2
2420882031742656142811241016325820316339691766342860 113712420753610 97417857829574741423886
471728262758155916515335 713174319356324651233421947 2 9452428483610 36615567239455348432250
067 0 922 118 053 8 918724150 8 6 2 1 8 316 3 216 17920 32817 3276317 2 0 1 1 616 3 5 229 0
1928491339603059 32015201017272323 9172767 93173 938 8 3631715474821 530 2351366524457504840
9522036371210 123 3 1 5171441252241 610 010 5 6 6 5442917334011 914383626 2 21530 3351417 0
012 0 820 025 025 6 0 5515324 526 3 0 0 136 7 010 73535 21915 0 33821 140 1 0 1 7 5 5 914 0
010 01112 0 7 310 5 1 2 9 5 2 319 5 7 3 126 1 830 91022 21619 1132710 123 3 0 5 611 5 3 9 0
18 015413224342331612016 717 81429167317 71623 918731417701541 8231435361750 624274544181436
14294114608540354453155615 945774324234545205076 839 220493128163923232744251560383851314461
1945 518 3 316 911201625654316 53639 8501044 3 921 96720 42526183463 0 222 1 3 554 9183425 0
6644624657635683476175593045535724587545745667615369374467534369534150644073675429704967571
521213391021 93216385934 3 5 72227316716512055 54723 2151913373310 2 964 227581313441541 765
6745 649375066586476583032686945116315423448226268415756695028493478605757 951624456403775
1 1 1 5 2 1 0 0 820 5 9 2 0 0 0 663 231 0 0 0 012 02111 01015 5 1 215 223 0 1 0 2 0 0 0 1 0
59316038040229533047493018 7334940056926552674104238 828453236443013366326635634385035302678
1638501323653740 520 917121432311515174770142689 84213 36015 6474726 535 6301663523357444843
82367338406638822560937210 83658355178488317675950591617583436755715128213669057637453654095
0 412 117 3 1 514 7 0 3 1 2 5 3 817 126 0 1 0 3 7 0 1 3 2 6 1 8 2 516 11011 3 210 9101715 1
23929 6725837 83518 011 4 7526526 7 819 5 82457 322 1164220 8 3 3 353314116 345371659 63745
83549355819 448452778256151460065857895088293626171 245744348853915788952929676650067555799

BIBLIOGRAPHY

- Abler, T.S. (1968) Traffic pattern and exhibit design: a study of learning in the museum. In S.F.deBorhegyi and I.A. Hanson (eds.) The Museum Visitor, Milwaukee Public Museum Publications in Museology, 3.
- Alt, M. B. (1979) Improving audio-visual presentations, Curator, 22, 2, 85 - 95.
- Alt, M. B. (1980) Four years of visitor surveys at the British Museum (Natural History) 1976-1979. Museums Journal, 80, 1, 10-19.
- Alt, M. B. (In Press) Designing and carrying out the evaluation study. In R. S. Miles (ed) The design of educational exhibits, London; George, Allen & Unwin.
- Alt, M. B. and Morris, R.G.M. (1979) The human biology exhibition at the Natural History Museum, Bulletin of the British Psychological Society, 32 273-78
- Bechtel, R.B. (1967) Hodometer research in museums, Museum News, 45, 7, 23-26
- Bloom, B.S., et al (1956) Taxonomy of educational objectives: handbook I: cognitive domain, New York: David McKay.
- Brandt, N. (1945) The psychology of seeing, New York: Philosophical Library.
- Bretz, R., (1971) A taxonomy of communication media, Englewood Cliffs, New Jersey: Educational Technology Pubs
- Bruner, J.S. (1960) The process of education, Cambridge: Harvard University Press.
- Buswell, G.T. (1935) How people look at pictures, Chicago: University of Chicago Press
- Cherry, C. (1961) On human communication, New York: Science Editions.

- Chevreul, M.E., (1967) The principle of harmony and contrast of colours. New York: Science Editions.
- Compton, A.H., (1935) The freedom of man. London, Oxford University Press.
- Craik, K.H. (1970) Environmental Psychology. In K.H. Craik and others, New directions in psychology, IV. New York: Holt, Rinehart and Winston.
- Craik, K.H. (1976) The personality research paradigm in environmental psychology. In S. Wapner, et al (eds.) Experiencing the environment. New York: Plenum.
- Cronbach, L.J., et al (1972) The dependability of behavioural measurements: theory of generalizability for scores and profiles. New York: Wiley
- Deutsch, M. & Krauss, R.M. (1965) Theories in social psychology. New York: Basic Books.
- De Waard, R.J. et al (1974) Effects of using programmed cards on learning in a museum environment. The Journal of educational research, 67, 10, 457-60.
- Dewey, J. (1963) Experience and education. New York: Crowell, Collier, Macmillan. (originally published by Kappa Delta Pi, 1938)
- Dwyer, F.M. (1971) Assessing students' perceptions of the instructional value of visual illustrations used to complement programmed instruction. Programmed learning and educational technology., 8, 73 - 80.
- Eason, L.P. & Linn, M.C. (1976) Evaluation of the effectiveness of participatory exhibits. Curator, 19 1, 45-62.
- Elliot, P. & Loomis, R.J. (1975) In A. Berman (ed.) Studies of visitor behaviour in museums and exhibitions: an annotated bibliography of sources primarily in the English language. Washington D.C.: Office of museum programs, Smithsonian Institute.
- Everitt, B (1974) Cluster analysis. London: Heinemann.
- Everitt, B. (1978) Graphical representations of multivariate data. London: Heinemann.

- Fransella, F. & Banister, D. . A manual for repertory grid technique,
(1977) London: Academic Press
- Gabriel, K.R. (1971) The biplot display of matrices with
application to principal components
analysis. Biometrika, 58, 3, 453-67
- Gagné, R.M. (1970) The conditions of learning, 2nd edition,
New York: Holt, Rinehart & Winston.
- Gagné, R.M. & Briggs, L.J. Principles of instructional design, New
(1974) York, Holt, Rinehart
- Gregory, R.L. (1966) Eye and brain - the psychology of seeing,
London: Weidenfeld & Nicolson.
- Gregory, R.L. (1978) Editorial, Perception, 7, 127.
- Gregory, R.L. (1980) Perceptions as hypotheses, Philosophical
transactions of the Royal Society of
London, B. biological sciences, 290,
1038, 181-98.
- Griggs, S.A. & Alt, M.B. (In Press) Visitors to the British
Museum (Natural History) in 1980 & 1981,
Museums Journal.
- Gauld, A. & Shotter, J. Human action and its psychological
(1977) investigation, London: Routledge & Kegan
Paul.
- Harré, R. (1972) The philosophies of science: an intro-
ductory survey, London: Oxford University
Press.
- Harré, R & Secord, P.F. The explanation of social behaviour.
(1972) Oxford: Blackwell.
- Herbert, J. & Attridge, C. A guide for developers and users of
(1975) observation systems and manuals. American
education research journal, 12, 1, 1-20.
- Hollenbeck, A.R. (1977) Problems of reliability in observational
research. In H. Sackett (ed.) Observing
behaviour, Baltimore: University Park
Press.

- Hull, D.L. (1968) . The operational imperative: sense and nonsense in operationalism. Systematic zoology, 17, 4, 438-57.
- Kerlinger, F.M. (1964) Foundations of behavioural research, 1st edition, New York: Holt, Rinehart & Winston.
- Kimsche, L. (1978) Science centers: a potential for learning. Science, 199, 270-73.
- Krathwohl, D.R. et al (1964) Taxonomy of educational objectives: hand-book II: affective domain, New York: David McKay.
- Kruskal, J.B. (1965) Analysis of factorial experiments by existing monotone transformations of the data. Journal of the Royal Statistical Society, 27, 251-63.
- Lakatos, I. (1964) Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (eds.) Criticism and the growth of knowledge, London: Cambridge University Press.
- Lakota, R.A. (1975) The National Museum of Natural History as a behavioural environment. Part I. An environmental analysis of behavioural performance. Washington D.C.: Office of programs, Smithsonian Institute.
- Lakota, R.A. & Kantner, J.A. (1976) The National Museum of Natural History as a behavioural environment. Part II. Summary and recommendations. Washington D.C. Office of museum programs, Smithsonian Institute.
- Loomis, R.J. (1973a) Museums and psychology: the principle of allometry and museum visitor research. The Museologist, 129, 17-23.
- Loomis, R.J. (1973b) Please! Not another visitor survey! Museum News, 52, 2, 21-26.
- MacBrier, W.N. (1964) Testing your audience, Museum News, 42, 8, 15-17.

- Mackworth, N. and Morandi, A. (1967) The gaze selects informative details within pictures, Perception and psychophysics, 2, (ii), 547-552
- Mager, R.F. (1962) Preparing instructional objectives. Belmont, California: Fearson.
- Maxwell, A.E. (1961) Analysing qualitative data, London: Methuen.
- Maxwell, A.E. and Pilliner, A.E.G.. (1968) Deriving coefficients of reliability and agreement between raters. British journal of mathematical and statistical psychology 21, 105-16.
- Maw, W.H. and Maw, E.W. (1978) Nature and assessment of human curiosity. In P. McReynolds (ed.) Advances in psychological assessment, volume 4, London: Jossey Bass.
- McNemar, Q. (1962) Psychological statistics, 3rd edition, New York: Wiley.
- Melton, .A.W (1933a) Studies of installation at the Pennsylvania Museum of Art, Museum news, 11, 5-8.
- Melton, A.W. (1933b) Some behavioural characteristics of museum visitors. Psychological bulletin, 30, 720-21.
- Melton, A.W. (1935) Problems of installation in museums of art, Publication of the American Association of Museums, New Series, No. 14.
- Melton, A.W. (1936) Distribution of attention in galleries in a museum of science and industry. Museum news, 14, 6-8.
- Melton, A.W. (1972) Visitor behaviour in museums: some early research in environmental design, Human factors, 14, 5, 393-403
- Melton, A.W. et al (1936) Experimental studies of the education of children in a museum of science. Publication of the American Association of Museums, New Series, No. 15.

- Miles, R.S. and Alt, M.B. (1979) British Museum (Natural History): a new approach to the visiting public. Museums Journal, 78, 4, 158-62.
- Miles, R.S. and Tout, A.F. (1978) Human biology and the new exhibition scheme in the British Museum (Natural History). Curator, 21, 1, 36-50
- Miles, R.S. and Tout, A.F. (1979) Outline of a technology for effective science exhibit. In Curation of palaeontological collections, special papers in palaeontology. 22, 209-224
- Mitchell, S.K. (1979) Interobserver agreement, reliability and generalizability of data collected in observational studies. Psychological bulletin. 86, 2, 376-90.
- Munyer, E.A. (1969) A preliminary survey of the Hall of Physical Anthropology, Museum of Natural History. Washington D.C.: Smithsonian Institution Summer Institute in Display.
- Nedzel, L.N. (1952) The motivation and education of the general public through museum experiences. Ph.D. dissertation: University of Chicago.
- Nielson, L.C. (1946) A technique for studying the behaviour of museum visitors. Journal of educational psychology, 37, 103-10
- O'Hare, D. (1979) Multidimensional scaling representations and the concept learning of artistic style. British journal of psychology, 70, 219-30.
- O'Hare, D. (1980) An introduction to the concepts and methods of multidimensional scaling. British educational research journal, 6, 1, 29-40.
- Oppenheimer, F. (1968) A rationale for a science museum. Curator, 11, 2, 206-09.
- Parsons, L.A. (1968) Systematic testing of display techniques for an anthropological exhibit. In S.F.de Borhegyi & I.A. Hanson (eds.) The museum visitor. Milwaukee Public Museum, Publications in Museology, 3.

- Parsons, M and Loomis, R.J. . Visitor traffic patterns: then and now. (1973) Washington D.C.: Office of Museums programs, Smithsonian Institute.
- Piaget, J. (1972) The principle of genetic epistemology. (Trans. by W. Mays.) London: Routledge and Kegan Paul.
- Porter, M.C.B. (1938) Behaviour of the average visitor in the Peabody Museum of Natural History, Yale University. Publications of the American Association of Museums, New Series, No. 16.
- Popper, K. R. (1972) Objective knowledge. An evolutionary approach. Oxford: Clarendon Press.
- Proshansky, H.M. (1970) In H.M. Proshansky et al (eds.) Environmental psychology. New York: Holt, Rinehart and Winston.
- Robinson, E.S. (1928) The behaviour of the Museum visitor. Publication of the American Association of Museums, New Series, No. 5.
- Robinson, E.S. (1930) Psychological problems of the science museum, Museum news, 8, 9-11.
- Robinson, E.S. (1931a) Psychological studies of the public museum. School and society, 33, 1-5.
- Robinson, E.S. (1931b) Exit the typical visitor. Journal of adult education, 3, 418-23.
- Robinson, E.S. (1933a) Psychology and public policy. School and Society, 37, 1-7.
- Robinson, E.S. (1933b) Experimental education in the museum - a perspective. Museum news, 11, 6-8.
- Rosch, E. (1978) Principles of categorisation. In E. Rosch and B. Lloyd (eds.) Cognition and categorisation. New Jersey: Erlbaum.
- Screven, C.G. (1968) The application of programmed learning and teaching systems procedures for instruction in a museum environment. In S.F. de Borhegyi and I.A. Hanson (eds.) The Museum Visitor. Milwaukee Public Museum Publications in Museology, 3.

- Screven, C.G. (1970) The programming of evaluation of an exhibit learning situation. In K.J. Goldman (ed.) Opportunities for extending museum contributions to pre-college science education. Washington D.C.: Smithsonian Institute informal publication.
- Screven, C.G. (1974) Instructional design, Museum news, 52, 5, 67-75.
- Screven, C.G. (1975) The measurement and facilitation of learning in the museum environment: an experimental analysis. Washington D.C.: Smithsonian Institution press.
- Screven, C.G. (1976) Exhibit evaluation: a goal referenced approach. Curator, 19, 4, 271-90
- Screven, C.G. and Brown, W. (1978) An evaluation case study at the British Columbia Provincial Museum. Milwaukee: Department of Psychology, University of Wisconsin.
- Screven C.G. and Lakota, R.A. (1970) An experimental study of learning in a museum environment. Proceedings of the annual convention of the American Psychological Association, 5, 2, 627-28.
- Scriven, M. (1967) The methodology of evaluation. In R.E. Stake (ed.) Perspectives in curriculum evaluation. American Educational Research Association Monograph Series on Evaluation, No. 1. Chicago: Rand McNally.
- Sheppard, D. (1960) Methods of assessing the value of exhibitions. British journal of educational psychology, 30, 259-65.
- Shettel, H.H. (1968) An evaluation of existing criteria for judging the quality of science exhibits. Curator, 11, 2, 137-53.
- Shettel, H.H. et al (1968) Strategies for determining exhibit effectiveness. Technical report No. AIR-E95-4/68-FR. Pittsburgh: American Institute for Research.
- Shettel, H.H. (1973) Exhibits: art form or educational medium. Museum news, 52, 1, 32-41.

- Siegel, S. (1956) Nonparametric statistics: for the behavioural sciences. International students edition. New York: McGraw-Hill.
- Sluckin, W. et al (1982) Some experimental studies of familiarity and liking. Bulletin of the British Psychological Society, 35, 189-94.
- Spencer, H. (1969) The visible word. London: Lund Humphries.
- Tversky, A. (1977) Features of similarity. Psychological Review, 84, 4, 327-52
- Tversky, A. and Gati, I. (1978) Studies in similarity. In E. Rosch and B. Lloyd (eds.) Cognition and categorisation. New Jersey: Erlbaum.
- Wallace, W. (1969) Sociological theory. London: Heinemann
- Webb, E.J. et al (1966) Unobtrusive measures: nonreactive research in the social sciences. Chicago: Rand McNally.
- Weiss, R.S. and Bourtourline, S. (1963) The communication value of exhibits. Museum news, 42, 3, 23-27.
- Wilson, J.T. (1976) Ontario Science Centre attracts the millions. Canadian Geographical Journal, March/April, 1-8.
- Winkel, G.H. and Sasanoff, R. (1970) An approach to an objective analysis of behaviour in architectural space. In H.M. Proshansky, et al (eds.) Environmental psychology. New York: Holt, Rinehart and Winston.
- Wittlin, A.S. (1971) Hazards of communication by exhibits. Curator, 14, 138-50.
- Yarbus, A.L. (1967) Eye movements and vision. New York: Plenum Press.